The Open University

M208

Pure mathematics

Handbook

This publication forms part of an Open University module. Details of this and other Open University modules can be obtained from Student Recruitment, The Open University, PO Box 197, Milton Keynes MK7 6BJ, United Kingdom (tel. +44 (0)300 303 5303; email general-enquiries@open.ac.uk).

Alternatively, you may visit the Open University website at www.open.ac.uk where you can learn more about the wide range of modules and packs offered at all levels by The Open University.

The Open University, Walton Hall, Milton Keynes, MK7 6AA.

First published 2019.

Edited, designed and typeset by The Open University, using LaTeX.

Printed in the United Kingdom by Hobbs the Printers Limited, Brunel Road, Totton, Hampshire, SO40 3WX.

SUP 05815 2

2.1

Contents

Notation

Standard symbols

$=$	is equal to
$\neq$	is not equal to
$\approx$	is approximately equal to
$<$	is less than
$\leq$	is less than or equal to
$>$	is greater than
$\geq$	is greater than or equal to
$\sqrt{\ }$	non-negative square root
∞	infinity
π	the irrational number $3.141\,59\ldots$
e	the irrational number $2.718\,28\ldots$ (also the identity element of a group, or the eccentricity of a conic)
$n!$	n factorial: $0! = 1$, $\quad n! = n \times (n-1) \times (n-2) \times \cdots \times 3 \times 2 \times 1 \quad$ for $n \in \mathbb{N}$
$\binom{n}{k}$	binomial coefficient: $\binom{n}{k} = \dfrac{n!}{k!\,(n-k)!}$, $\quad$ for $k \leq n$ and $k, n \in \mathbb{N}$

Standard sets of numbers

$\mathbb{R}$	set of real numbers
$\mathbb{Q}$	set of rational numbers
$\mathbb{Z}$	set of integers, $\{\ldots, -2, -1, 0, 1, 2, \ldots\}$
$\mathbb{N}$	set of natural numbers, $\{1, 2, 3, \ldots\}$
$\mathbb{C}$	set of complex numbers
$\mathbb{Z}_n$	set of integers modulo n, $\{0, 1, \ldots, n-1\}$
$n\mathbb{Z}$	set of integer multiples of number n
$\mathbb{R}^+$	set of positive real numbers
$\mathbb{Q}^+$	set of positive rational numbers
$\mathbb{Z}^+$	set of positive integers (equal to $\mathbb{N}$)
$\mathbb{R}^*$	set of non-zero real numbers
$\mathbb{Q}^*$	set of non-zero rational numbers
$\mathbb{Z}^*$	set of non-zero integers
$\mathbb{C}^*$	set of non-zero complex numbers
$\mathbb{Z}_n^*$	set of non-zero integers modulo n, $\{1, 2, 3, \ldots, n-1\}$
$\mathbb{R}^2$	set of all ordered pairs of real numbers, also called the plane
$\mathbb{R}^3$	set of all ordered triples of real numbers, also called three-dimensional space
$\mathbb{R}^n$	set of all ordered n-tuples of real numbers, called n-dimensional space
$\mathbb{R}^\infty$	set of all infinite sequences of real numbers, $\{(v_1, v_2, v_3, \ldots) : v_i \in \mathbb{R}\}$

Greek alphabet

α	A	alpha	η	H	eta	ν	N	nu	τ	T	tau
β	B	beta	θ	Θ	theta	ξ	Ξ	xi	υ	Υ	upsilon
γ	Γ	gamma	ι	I	iota	o	O	omicron	ϕ	Φ	phi
δ	Δ	delta	κ	K	kappa	π	Π	pi	χ	X	chi
ε	E	epsilon	λ	Λ	lambda	ρ	P	rho	ψ	Ψ	psi
ζ	Z	zeta	μ	M	mu	σ	Σ	sigma	ω	Ω	omega

Notation introduced in Introduction (Book A)

$\{x, y, \dots, z\}$	set of elements listed in $\{\dots\}$
$\{x : \dots\}$	set of all x such that ...
(a, b)	open interval, excluding endpoints a, b, $\{x : a < x < b\}$
$[a, b]$	closed interval, including endpoints a, b, $\{x : a \leq x \leq b\}$
$(a, b]$	half-open interval, excluding a, including b, $\{x : a < x \leq b\}$
$[a, b)$	half-open interval, including a, excluding b, $\{x : a \leq x < b\}$
$(-\infty, a)$	open interval, $\{x : x < a\}$
$(-\infty, a]$	closed interval, $\{x : x \leq a\}$
(a, ∞)	open interval, $\{x : x > a\}$
$[a, \infty)$	closed interval, $\{x : x \geq a\}$
$\varnothing$	empty set
$a \in A$	a is an element of set A
$a \notin A$	a is not an element of set A
$A \subseteq B$	set A is a subset of set B
$A \subset B$	set A is a proper subset of set B
$A = B$	set A is equal to set B, $A \subseteq B$ and $B \subseteq A$
$A \cup B$	union of sets A and B, $\{x : x \in A \text{ or } x \in B\}$
$A \cap B$	intersection of sets A and B, $\{x : x \in A \text{ and } x \in B\}$
$A - B$	difference between sets A and B, $\{x : x \in A,\ x \notin B\}$
$\lvert x \rvert$	modulus of real number x
$\lfloor x \rfloor$	integer part of real number x
$\longrightarrow$	maps to, for sets
$\longmapsto$	maps to, for variables
$g \circ f$	composite function with rule $x \longmapsto g(f(x))$, where f and g are functions
f^{-1}	inverse of function f
$\sin^{-1}$	inverse of function sin (similarly for cos, tan, sinh, etc.)
exp	*the* exponential function $f(x) = e^x$
$f'(x), f''(x)$	first and second derivatives of function f at x
$\to$	tends to (for asymptotic behaviour and limits)
$\lvert \mathbf{v} \rvert$	magnitude of vector $\mathbf{v}$
$\mathbf{0}$	zero vector (or zero matrix)
$\mathbf{i}, \mathbf{j}, \mathbf{k}$	unit vectors in the directions of the x-, y-, z-axes, respectively
(a, b)	a vector in $\mathbb{R}^2$ written in component form
(a, b, c)	a vector in $\mathbb{R}^3$ written in component form

$\mathbf{u} \cdot \mathbf{v}$	scalar product of vectors $\mathbf{u}$ and $\mathbf{v}$
$\widehat{\mathbf{v}}$	unit vector in the same direction as vector $\mathbf{v}$
$\overrightarrow{AB}$	a vector represented by a line segment from A to B
$x + iy$	a complex number, where x and y are real numbers and $i^2 = -1$
$\overline{z}$	complex conjugate of complex number z
$\lvert z \rvert$	modulus of complex number z
$\operatorname{Re} z$	real part of complex number z
$\operatorname{Im} z$	imaginary part of complex number z
$\arg z$	an argument of complex number z
$\operatorname{Arg} z$	principal argument of complex number z
$P \implies Q$	if P, then Q (P implies Q)
$P \iff Q$	P if and only if Q (P is equivalent to Q)
$\forall$	the universal quantifier, 'for all'
$\exists$	the existential quantifier, 'there exists'
$a \equiv b \pmod{n}$	a is congruent to b modulo n
$a +_n b$	remainder of $a + b$ on division by n
$a \times_n b$	remainder of $a \times b$ on division by n
$x \sim y$	x is related to y (by a particular relation)
$x \not\sim y$	x is not related to y (by a particular relation)
$[\![x]\!]$	equivalence class of x (with respect to a particular equivalence relation)
e	eccentricity of a conic

Notation introduced in Linear algebra (Book C)

$\mathbf{r}_i \leftrightarrow \mathbf{r}_j$	interchange r_i and r_j
$\mathbf{r}_i \to \alpha\mathbf{r}_i$	multiply r_i by α (non-zero)
$\mathbf{r}_i \to \mathbf{r}_i + \beta\mathbf{r}_j$	change r_i to r_i plus βr_j
(a_{ij})	matrix with (i, j)-entry a_{ij}
$(\mathbf{A} \mid \mathbf{b})$	augmented matrix of system $\mathbf{Ax} = \mathbf{b}$ of linear equations
$(\mathbf{A} \mid \mathbf{I})$	matrix $\mathbf{A}$ augmented by matrix $\mathbf{I}$
$\mathbf{0}_{m,n}$	zero matrix of size $m \times n$
$\mathbf{I}_n$	identity matrix of size $n \times n$
$\mathbf{A}^T$	transpose of matrix $\mathbf{A}$
$\mathbf{A}^{-1}$	inverse of matrix $\mathbf{A}$
$\det \mathbf{A}$	determinant of matrix $\mathbf{A}$
A_{ij}	cofactor associated with entry a_{ij} of matrix $\mathbf{A}$
$M_{m,n}$	set of all $m \times n$ matrices with real entries
P_n	set of all real polynomials of degree less than n
$\langle S \rangle$	span of finite set S of vectors
$\mathbf{v}_E$	E-coordinate representation of vector $\mathbf{v}$ (coordinates with respect to basis E)
$\dim V$	dimension of vector space V
i_V	identity linear transformation of vector space V
$\operatorname{Im} t$	image set of linear transformation t
$\operatorname{Ker} t$	kernel of linear transformation t
$S(\lambda)$	eigenspace corresponding to eigenvalue λ

Notation introduced in Group theory (Books B and E)

$S(F)$	set of all symmetries of figure F, a group under function composition
$S^+(F)$	set of all direct symmetries of figure F, a group under function composition
$S(\triangle)$	group of all symmetries of the equilateral triangle
$S(\square)$	group of all symmetries of the square
$S(▭)$	group of all symmetries of the rectangle
$S(✣)$	group of all symmetries of the 4-windmill
$S(\bigcirc)$	(infinite) group of all symmetries of the disc
r_θ	rotation through θ (anticlockwise) about the centre of a disc (or the origin)
q_θ	reflection in the line through the centre of a disc (or the origin) at angle θ (anticlockwise) to the horizontal
$(G, \circ)$	set G with binary operation $\circ$
$\lvert G\rvert$	order of group G
$\langle x\rangle$	cyclic group generated by x
e	identity element of a group
x^{-1}	inverse of group element x, in multiplicative notation
$-x$	inverse of group element x, in additive notation
$\cong$	is isomorphic to
V	Klein four-group
C_n	standard abstract cyclic group of order n (*the* cyclic group of order n)
U_n	set of all integers in $\mathbb{Z}_n$ coprime to n, a group under $\times_n$
S_n	symmetric group of degree n (of order $n!$)
A_n	alternating group of degree n (of order $n!/2$ for $n \geq 2$, and order 1 for $n = 1$)
D_n	dihedral group of order $2n$ (the symmetry group of an n-gon), $n \geq 3$
Q_8	quaternion group of order 8
$\mathrm{GL}(2)$	general linear group of degree 2 (invertible 2×2 matrices under matrix multiplication)
$\mathrm{SL}(2)$	special linear group of degree 2 (2×2 matrices with determinant 1 under matrix multiplication)
D	group of invertible 2×2 diagonal matrices under matrix multiplication
U	group of invertible 2×2 upper triangular matrices under matrix multiplication
L	group of invertible 2×2 lower triangular matrices under matrix multiplication
$\mathrm{GL}(n)$	general linear group of degree n (invertible $n \times n$ matrices under matrix multiplication)
$\mathrm{SL}(n)$	special linear group of degree n ($n \times n$ matrices with determinant 1 under matrix multiplication)
$\mathrm{frac}(x)$	fractional part of real number x
gH, Hg	left coset and right coset of subgroup H in a particular group
$A \cdot B$	composite, under set composition, of subsets A, B of a group $(G, \circ)$, $\{a \circ b : a \in A,\ b \in B\}$
G/N	quotient group of group G by normal subgroup N
gHg^{-1}	conjugate subgroup of subgroup H by group element g
$k_{(\mathrm{mod}\ n)}$	least residue of k modulo n, the reminder of k on division by n
$\mathrm{Im}\,\phi$	image of homomorphism ϕ
$\mathrm{Ker}\,\phi$	kernel of homomorphism ϕ
$g \wedge x$	set element obtained when group element g acts on set element x
$\mathrm{Orb}\,x$	orbit of x
$\mathrm{Stab}\,x$	stabiliser of x
$\mathrm{Fix}\,g$	fixed set (or fixed point set) of g

Notation introduced in Analysis (Books D and F)

$\max E$	maximum element of set E
$\min E$	minimum element of set E
$\sup E$	supremum (least upper bound) of set E
$\inf E$	infimum (greatest lower bound) of set E
$\sqrt[n]{a}$	nth root of non-negative real number a
(a_n)	sequence of numbers, $a_1, a_2, \ldots$
$\lim\limits_{n\to\infty} a_n$	limit of sequence (a_n) as n tends to ∞
s_n	nth partial sum, $a_1 + a_2 + \ldots + a_n$
$\sum\limits_{n=1}^{\infty} a_n$	sum of series $a_1 + a_2 + \cdots$
$\not\to$	does not tend to
$N_r(c)$	punctured neighbourhood of c, $(c-r,c)\cup(c,c+r)$, $r>0$
$\lim\limits_{x\to c}$	limit as x tends to c
$\lim\limits_{x\to c^+}$	limit as x tends to c from the right
$\lim\limits_{x\to c^-}$	limit as x tends to c from the left
$Q(h)$	difference quotient, the gradient of the line through $(c, f(c))$ and $((c+h), f(c+h))$
$\dfrac{dy}{dx}, \dfrac{d^2y}{dx^2}$	Leibniz notation for first and second derivatives of $y = f(x)$ with respect to x
$f^{(n)}$	nth derivative of function f
$f'_L(c)$	left derivative of f at c
$f'_R(c)$	right derivative of f at c
$\min\limits_{[a,b]} f$	minimum of function f on $[a,b]$
$\max\limits_{[a,b]} f$	maximum of function f on $[a,b]$
$\inf\limits_{[a,b]} f$	infimum of function f on $[a,b]$
$\sup\limits_{[a,b]} f$	supremum of function f on $[a,b]$

Notation

P	partition of an interval
δx_i	length of the ith subinterval $[x_{i-1}, x_i]$ of a partition
$\|P\|$	mesh of partition P, length of longest subinterval of P
m_i	$\inf\{f(x) : x_{i-1} \le x \le x_i\}$, where f is a function on $[x_{i-1}, x_i]$
M_i	$\sup\{f(x) : x_{i-1} \le x \le x_i\}$, where f is a function on $[x_{i-1}, x_i]$
$L(f, P)$	lower Riemann sum of function f over partition P
$U(f, P)$	upper Riemann sum of function f over partition P
$\underline{\int_a^b} f$	lower integral of function f over interval $[a, b]$
$\overline{\int_a^b} f$	upper integral of function f over interval $[a, b]$
$\int_a^b f$	integral of function f over interval $[a, b]$
$\int f(x)\,dx$	a primitive of function f, often denoted by F
$\big[F(x)\big]_a^b$	primitive F evaluated between a and b, $F(b) - F(a)$
$\sim$	$f(n) \sim g(n)$ as $n \to \infty$ means $f(n)/g(n) \to 1$ as $n \to \infty$
$T_n(x)$	Taylor polynomial of degree n
$R_n(x)$	remainder term for Taylor polynomial of degree n
R	radius of convergence of a power series
$\binom{\alpha}{n}$	generalised binomial coefficient, $\alpha \in \mathbb{R}$: $\binom{\alpha}{0} = 1, \quad \binom{\alpha}{n} = \dfrac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!} \quad \text{for } n \in \mathbb{N}$

Book A Introduction

Unit A1 Sets, functions and vectors

1 Points, lines and distance

1. Each point in the plane is represented by an ordered pair (x, y) of real numbers. The plane $\mathbb{R}^2$, together with an origin and a pair of x- and y-axes, is known as **two-dimensional Euclidean space**.

2. Equation of a line The general equation of a line in $\mathbb{R}^2$ is

$$ax + by = c,$$

where a, b and c are real numbers, and a and b are not both zero.

Two distinct lines are **parallel** if they never meet, and **perpendicular** if they meet at right angles.

Two non-vertical lines with gradients m_1 and m_2, respectively, are perpendicular if and only if $m_1 m_2 = -1$.

3. The **modulus** of a real number k is

$$|k| = \begin{cases} k, & \text{if } k \geq 0, \\ -k, & \text{if } k < 0. \end{cases}$$

4. Distance formula for $\mathbb{R}^2$ The distance between the points (x_1, y_1) and (x_2, y_2) in $\mathbb{R}^2$ is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}.$$

5. Equation of a circle The equation of the circle in $\mathbb{R}^2$ with centre (a, b) and radius r is

$$(x - a)^2 + (y - b)^2 = r^2.$$

6. Each point in three-dimensional space is represented by an ordered triple (x, y, z) of real numbers. Three-dimensional space $\mathbb{R}^3$, together with an origin and a set of x-, y- and z-axes, is known as **three-dimensional Euclidean space**.

7. Distance formula for $\mathbb{R}^3$ The distance between the points (x_1, y_1, z_1) and (x_2, y_2, z_2) in $\mathbb{R}^3$ is

$$\sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}.$$

2 Sets

8. A **set** is a collection of objects, such as numbers, points, functions, or even other sets. Each object in a set is an **element** or **member** of the set, and the elements *belong to* the set, or are *in* the set.

9. A **finite set** is a set that has a finite number of elements; that is, the number of elements is some natural number, or 0. An **infinite set** is any set that is not a finite set.

A **singleton** is a set with only one element, such as the set $\{2\}$.

The **empty set**, denoted by $\varnothing$, has no elements.

10. To indicate that a is an element of the set A, we write $a \in A$.

To indicate that b is not an element of the set A, we write $b \notin A$.

11. The **solution set** of an equation, or a system of equations, is the set of its solutions. It depends on the set from which the solutions are taken.

12. An **interval** is a set comprising a range of real numbers.

Intervals are denoted as follows.

Open intervals

(a, b) — $a < x < b$

(a, ∞) — $x > a$

$(-\infty, b)$ — $x < b$

Closed intervals

$[a, b]$ — $a \leq x \leq b$

$[a, \infty)$ — $x \geq a$

$(-\infty, b]$ — $x \leq b$

Open and closed interval

$(-\infty, \infty)$ — $\mathbb{R}$

Half-open (or half-closed) intervals

$[a, b)$ — $a \leq x < b$

$(a, b]$ — $a < x \leq b$

13. A **plane set** (or **plane figure**) is a set of points in $\mathbb{R}^2$. Simple examples of plane sets are lines and circles.

14. The **line** l with gradient m and y-intercept c is

$$\begin{aligned} l &= \{(x, y) \in \mathbb{R}^2 : y = mx + c\} \\ &= \{(x, mx + c) : x \in \mathbb{R}\}. \end{aligned}$$

We refer to 'the line $y = mx + c$' as a shorthand way of specifying this set.

A **half-plane** is the set of points on one side of a line, possibly together with all the points on the line itself.

15. The **circle** C with centre (a, b) and radius r is

$$C = \{(x, y) \in \mathbb{R}^2 : (x - a)^2 + (y - b)^2 = r^2\}.$$

The **unit circle** U is

$$\{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}.$$

A **disc** is the set of points inside a circle, possibly together with all the points on the circle.

16. Convention for drawing sets in $\mathbb{R}$ or $\mathbb{R}^2$

- included and excluded points are drawn as solid and hollow dots, respectively
- included and excluded boundaries are drawn as solid and broken lines, respectively.

17. Two sets A and B are **equal** if they have exactly the same elements; we write $A = B$.

18. A **subset** of a set B is a set A such that each element of A is also an element of B. We say that A is *contained in* B or B *contains* A, and we write $A \subseteq B$ or $B \supseteq A$.

To indicate that A is *not* a subset of B, we write $A \nsubseteq B$ or $B \nsupseteq A$.

The subsets of a set B include the empty set $\varnothing$ and the set B itself.

A **proper subset** of a set B is a set A that is a subset of B, but is not equal to B; we write $A \subset B$ or $B \supset A$.

19. Showing that two sets are equal

Strategy A1 Set equality

To show that the sets A and B are equal:

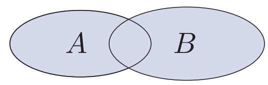
- first show that $A \subseteq B$
- then show that $B \subseteq A$.

20. Set operations

Let A and B be any sets.

The **union** of A and B is the set

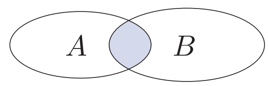
$$A \cup B = \{x : x \in A \text{ or } x \in B\}.$$



$A \cup B$

The **intersection** of A and B is the set

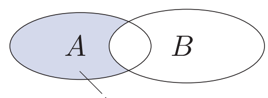
$$A \cap B = \{x : x \in A \text{ and } x \in B\}.$$



$A \cap B$

The **difference** between A and B is the set

$$A - B = \{x : x \in A,\ x \notin B\}.$$



$A - B$

Disjoint sets have no elements in common.

3 Functions

21. Functions

A **function** f is defined by specifying:

- a set A, called the **domain** of f
- a set B, called the **codomain** of f
- a **rule** $x \longmapsto f(x)$ that associates each element $x \in A$ with a unique element $f(x) \in B$.

The element $f(x)$ is the **image** of x under f.

Symbolically, we write

$$\begin{aligned} f : A &\longrightarrow B \\ x &\longmapsto f(x). \end{aligned}$$

We often refer to a function as a **mapping**, and say that f **maps** A to B and x to $f(x)$.

22. A **real function** is a function whose domain and codomain are both subsets of $\mathbb{R}$.

23. The **identity function** on a set A is the function

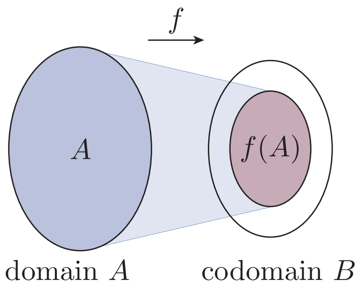
$$\begin{aligned} i_A : A &\longrightarrow A \\ x &\longmapsto x. \end{aligned}$$

24. Let $f : A \longrightarrow B$ be a function. For any subset S of A, the **image** of S under f, denoted by $f(S)$, is the set

$$f(S) = \{f(x) : x \in S\}.$$

The **image set** of a function $f : A \longrightarrow B$ is the set

$$f(A) = \{f(x) : x \in A\}.$$



25. A function $f : A \longrightarrow B$ is **onto** if $f(A) = B$.

26. A function $f : A \longrightarrow B$ is **one-to-one** if each element of $f(A)$ is the image of exactly one element of A; that is,

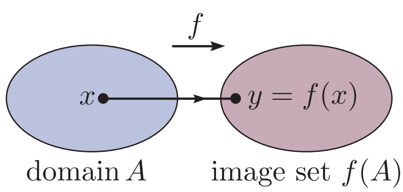
$$\text{if } x_1, x_2 \in A \text{ and } f(x_1) = f(x_2), \text{ then } x_1 = x_2.$$

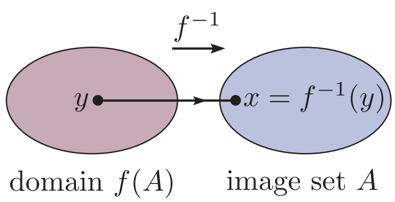
A function that is not one-to-one is **many-to-one**.

27. Inverse functions

Let $f : A \longrightarrow B$ be a one-to-one function. Then f has an **inverse function** $f^{-1} : f(A) \longrightarrow A$, with rule

$$f^{-1}(y) = x, \quad \text{where } y = f(x).$$





28. If a function $f : A \longrightarrow B$ is *onto* as well as one-to-one, then f has an inverse function $f^{-1} : B \longrightarrow A$; the functions f and f^{-1} are inverses of each other.

Such a function is a **one-to-one correspondence** between the sets A and B.

29. Let $f : A \longrightarrow B$ be a function and let C be a subset of the domain A. Then the **restriction** of f to C is the function $g : C \longrightarrow B$ defined by

$$g(x) = f(x), \quad \text{for } x \in C.$$

30. Composite functions

A **composite function** is a function obtained by applying first one function and then another.

Let $f : A \longrightarrow B$ and $g : B \longrightarrow C$ be functions such that the domain B of g is the same set as the codomain of f. Then the composite function $g \circ f$ (meaning 'f then g') is given by

$$\begin{aligned} g \circ f : A &\longrightarrow C \\ x &\longmapsto g(f(x)). \end{aligned}$$

More generally, let $f : A \longrightarrow B$ and $g : C \longrightarrow D$ be any functions; then the composite function $g \circ f$ is given by

$$\begin{aligned} g \circ f : A' &\longrightarrow D \\ x &\longmapsto g(f(x)), \end{aligned}$$

where $A' = \{x \in A : f(x) \in C\}$. Thus $g \circ f$ has

- domain $A' = \{x \in A : f(x) \in C\}$
- codomain D
- rule $(g \circ f)(x) = g(f(x))$.

31. Showing that a function is the inverse of another function

Strategy A2 Inverse functions

To show that the function $g : B \longrightarrow A$ is the inverse function of the function $f : A \longrightarrow B$:

1. show that $g(f(x)) = x$ for each $x \in A$; that is, $g \circ f = i_A$
2. show that $f(g(y)) = y$ for each $y \in B$; that is, $f \circ g = i_B$.

4 Vectors

32. A **vector** is a quantity that is determined by its magnitude and direction. A **scalar** is a quantity that is determined by its magnitude.

The **magnitude** of a vector $\mathbf{v}$ is denoted by $|\mathbf{v}|$.

33. The **zero vector** is the vector whose magnitude is zero, and whose direction is undefined. It is denoted by $\mathbf{0}$.

34. Two vectors $\mathbf{a}$ and $\mathbf{b}$ are **equal** if they have the same magnitude ($|\mathbf{a}| = |\mathbf{b}|$) and if they are in the same direction. We write $\mathbf{a} = \mathbf{b}$.

35. The **negative** of a vector $\mathbf{v}$ is the vector that has the same magnitude as $\mathbf{v}$, but the opposite direction. It is denoted by $-\mathbf{v}$.

36. Scalar multiple of a vector

Let k be a scalar and $\mathbf{v}$ a vector. The **scalar multiple** $k\mathbf{v}$ of $\mathbf{v}$ is the vector

- whose magnitude is $|k|$ times the magnitude of $\mathbf{v}$; that is, $|k\mathbf{v}| = |k|\,|\mathbf{v}|$
- that has the same direction as $\mathbf{v}$ if $k > 0$, and the opposite direction if $k < 0$.

If $k = 0$, then $k\mathbf{v} = \mathbf{0}$.

37. Addition of vectors

Triangle Law for addition of vectors
The sum $\mathbf{p}+\mathbf{q}$ of two vectors $\mathbf{p}$ and $\mathbf{q}$ is obtained as follows.

1. Starting at any point, draw the vector $\mathbf{p}$.
2. Starting from the tip of the vector $\mathbf{p}$, draw the vector $\mathbf{q}$.

Then the sum $\mathbf{p}+\mathbf{q}$ is the vector from the tail of $\mathbf{p}$ to the tip of $\mathbf{q}$.

Parallelogram Law for addition of vectors
The sum $\mathbf{p}+\mathbf{q}$ of two vectors $\mathbf{p}$ and $\mathbf{q}$ is obtained as follows.

1. Starting at the same point, draw the vectors $\mathbf{p}$ and $\mathbf{q}$.
2. Complete the parallelogram of which these vectors are adjacent sides.

Then the sum $\mathbf{p}+\mathbf{q}$ is the vector from the point where the tails of $\mathbf{p}$ and $\mathbf{q}$ meet to the opposite corner of the parallelogram.

Properties of vector algebra

Let $\mathbf{p}$, $\mathbf{q}$ and $\mathbf{r}$ be vectors, and let $a, b \in \mathbb{R}$. The following properties hold.

Commutativity $\quad \mathbf{p}+\mathbf{q} = \mathbf{q}+\mathbf{p}$

Associativity $\quad (\mathbf{p}+\mathbf{q})+\mathbf{r} = \mathbf{p}+(\mathbf{q}+\mathbf{r})$

Distributivity $\quad a(\mathbf{p}+\mathbf{q}) = a\mathbf{p}+a\mathbf{q}$,
$(a+b)\mathbf{p} = a\mathbf{p}+b\mathbf{p}$.

38. The **difference** $\mathbf{p}-\mathbf{q}$ of the vectors $\mathbf{p}$ and $\mathbf{q}$ is

$$\mathbf{p}-\mathbf{q} = \mathbf{p}+(-\mathbf{q}).$$

39. A **unit vector** is a vector of magnitude 1.

In $\mathbb{R}^2$, the vectors $\mathbf{i}$ and $\mathbf{j}$ are the unit vectors in the positive directions of the x- and y-axes, respectively.

In $\mathbb{R}^3$, the vectors $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are the unit vectors in the positive directions of the x-, y- and z-axes, respectively.

The unit vector in the same direction as a non-zero vector $\mathbf{v}$ is

$$\widehat{\mathbf{v}} = \frac{1}{|\mathbf{v}|}\mathbf{v}.$$

40. The **component form** of a vector $\mathbf{p}$ in $\mathbb{R}^2$ is

$$\mathbf{p} = a_1\mathbf{i}+a_2\mathbf{j}, \quad \text{where } a_1, a_2 \in \mathbb{R},$$

often written as $\mathbf{p} = (a_1, a_2)$.

The **components** of $\mathbf{p}$ in the x- and y-directions are the numbers a_1 and a_2, respectively.

The **component form** of a vector $\mathbf{p}$ in $\mathbb{R}^3$ is

$$\mathbf{p} = a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}, \quad \text{where } a_1, a_2, a_3 \in \mathbb{R},$$

often written as $\mathbf{p} = (a_1, a_2, a_3)$.

The **components** of $\mathbf{p}$ in the x-, y- and z-directions are the numbers a_1, a_2 and a_3, respectively.

41. Vector arithmetic in component form

Equality Two vectors, both in $\mathbb{R}^2$ or both in $\mathbb{R}^3$, are equal if their corresponding components are equal.

Zero vector The zero vector is

$$\mathbf{0} = (0,0) \quad \text{in } \mathbb{R}^2,$$

$$\mathbf{0} = (0,0,0) \quad \text{in } \mathbb{R}^3.$$

Addition To add vectors in $\mathbb{R}^2$ or in $\mathbb{R}^3$, add their corresponding components:

$$(a_1, a_2) + (b_1, b_2) = (a_1 + b_1, a_2 + b_2),$$
$$(a_1, a_2, a_3) + (b_1, b_2, b_3) = (a_1 + b_1, a_2 + b_2, a_3 + b_3).$$

Negatives To find the negative of a vector in $\mathbb{R}^2$ or in $\mathbb{R}^3$, take the negatives of its components:

$$-(a_1, a_2) = (-a_1, -a_2),$$
$$-(a_1, a_2, a_3) = (-a_1, -a_2, -a_3).$$

Subtraction To subtract vectors in $\mathbb{R}^2$ or in $\mathbb{R}^3$, subtract the corresponding components:

$$(a_1, a_2) - (b_1, b_2) = (a_1 - b_1, a_2 - b_2),$$
$$(a_1, a_2, a_3) - (b_1, b_2, b_3) = (a_1 - b_1, a_2 - b_2, a_3 - b_3).$$

Scalar multiplication To multiply a vector in $\mathbb{R}^2$ or in $\mathbb{R}^3$ by a real number k, multiply its components by k:

$$k(a_1, a_2) = (ka_1, ka_2),$$
$$k(a_1, a_2, a_3) = (ka_1, ka_2, ka_3).$$

Magnitude The magnitudes of the vectors (a_1, a_2) in $\mathbb{R}^2$ and (a_1, a_2, a_3) in $\mathbb{R}^3$ are

$$\sqrt{a_1^2 + a_2^2} \quad \text{and} \quad \sqrt{a_1^2 + a_2^2 + a_3^2},$$

respectively.

42. The **position vector** of a point P in $\mathbb{R}^2$ or $\mathbb{R}^3$ is the vector whose starting point is the origin O and whose finishing point is P; that is, the vector $\overrightarrow{OP}$.

Let A and B be points (in $\mathbb{R}^2$ or $\mathbb{R}^3$), with position vectors $\mathbf{a}$ and $\mathbf{b}$, respectively. Then

$$\overrightarrow{AB} = \mathbf{b} - \mathbf{a}.$$

43. Vector form of the equation of a line

The equation of the line through the points with position vectors $\mathbf{p}$ and $\mathbf{q}$ is

$$\mathbf{r} = (1 - \lambda)\mathbf{p} + \lambda\mathbf{q}, \quad \text{where } \lambda \in \mathbb{R}.$$

44. The **scalar product** (or **dot product**) of two non-zero vectors $\mathbf{u}$ and $\mathbf{v}$, in $\mathbb{R}^2$ or $\mathbb{R}^3$, is

$$\mathbf{u} \cdot \mathbf{v} = |\mathbf{u}||\mathbf{v}|\cos\theta,$$

where θ is the angle between $\mathbf{u}$ and $\mathbf{v}$.

If one or both of $\mathbf{u}$ and $\mathbf{v}$ is the zero vector, then $\mathbf{u} \cdot \mathbf{v} = 0$.

Algebraic properties of scalar product

Let $\mathbf{u}$, $\mathbf{v}$ and $\mathbf{w}$ be vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$, and let $\alpha \in \mathbb{R}$. The following properties hold.

Commutativity $\quad \mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$

Multiples $\quad (\alpha\mathbf{u}) \cdot \mathbf{v} = \mathbf{u} \cdot (\alpha\mathbf{v}) = \alpha(\mathbf{u} \cdot \mathbf{v})$

Distributivity $\quad \mathbf{u} \cdot (\mathbf{v} + \mathbf{w}) = \mathbf{u} \cdot \mathbf{v} + \mathbf{u} \cdot \mathbf{w},$
$(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}.$

45. Scalar product in component form

In $\mathbb{R}^2$, let $\mathbf{u} = (x_1, y_1)$ and $\mathbf{v} = (x_2, y_2)$. Then

$$\mathbf{u} \cdot \mathbf{v} = x_1x_2 + y_1y_2.$$

In $\mathbb{R}^3$, let $\mathbf{u} = (x_1, y_1, z_1)$ and $\mathbf{v} = (x_2, y_2, z_2)$. Then

$$\mathbf{u} \cdot \mathbf{v} = x_1x_2 + y_1y_2 + z_1z_2.$$

46. The angle θ between two vectors $\mathbf{u}$ and $\mathbf{v}$ is given by

$$\cos\theta = \frac{\mathbf{u} \cdot \mathbf{v}}{|\mathbf{u}||\mathbf{v}|}.$$

47. For any vector $\mathbf{v}$ in $\mathbb{R}^2$ or $\mathbb{R}^3$,

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}}.$$

48. Scalar product and perpendicularity

Let $\mathbf{u}$ and $\mathbf{v}$ be vectors in $\mathbb{R}^2$ or $\mathbb{R}^3$.

- If $\mathbf{u}$ and $\mathbf{v}$ are perpendicular, then $\mathbf{u} \cdot \mathbf{v} = 0$.
- If $\mathbf{u} \cdot \mathbf{v} = 0$, then $\mathbf{u} = \mathbf{0}$, or $\mathbf{v} = \mathbf{0}$, or $\mathbf{u}$ and $\mathbf{v}$ are perpendicular.

49. A **normal vector** (or simply a **normal**) to a plane is a vector that is perpendicular to all the vectors in the plane. Its direction is **normal** to the plane.

50. Equation of a plane in $\mathbb{R}^3$

The equation of the plane that contains the point (x_1, y_1, z_1) and has the vector $\mathbf{n} = (a, b, c)$ as a normal is

$$ax + by + cz = d,$$

where $d = ax_1 + by_1 + cz_1$.

This equation can be written in vector form as

$$\mathbf{x} \cdot \mathbf{n} = \mathbf{p} \cdot \mathbf{n},$$

where $\mathbf{x} = (x, y, z)$ and $\mathbf{p} = (x_1, y_1, z_1)$.

Unit A2 Number systems

1 Real numbers

1. The set of all **real numbers** is denoted by $\mathbb{R}$. It can be pictured as a number line, often called the **real line**. Each real number is represented by a point on the real line, and each point on this line represents a real number.

2. The set of all **rational numbers** is denoted by $\mathbb{Q}$. A **rational number** is a real number that can be expressed as a fraction whose numerator and denominator are integers:

$$\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{N}\}.$$

The set $\mathbb{Q}$ is a proper subset of $\mathbb{R}$. The real numbers that are not rational numbers are **irrational** numbers.

Theorem A1

There is no rational number x such that

$$x^2 = 2.$$

3. Every real number has a decimal expansion.

- The decimal expansion of a *rational* number is always either a terminating (or finite) decimal or a recurring decimal.
- The decimal expansion of an *irrational* number is neither terminating nor recurring: it continues for ever, with no repeating pattern of digits.

4. **Arithmetic in $\mathbb{R}$**

Properties for addition in $\mathbb{R}$

A1 Closure For all $a, b \in \mathbb{R}$, $a + b \in \mathbb{R}$.

A2 Associativity For all $a, b, c \in \mathbb{R}$,

$$a + (b + c) = (a + b) + c.$$

A3 Additive identity For all $a \in \mathbb{R}$,

$$a + 0 = a = 0 + a.$$

A4 Additive inverses For each $a \in \mathbb{R}$, there is a number $-a \in \mathbb{R}$ such that

$$a + (-a) = 0 = (-a) + a.$$

A5 Commutativity For all $a, b \in \mathbb{R}$,

$$a + b = b + a.$$

Properties for multiplication in $\mathbb{R}$

M1 Closure For all $a, b \in \mathbb{R}$, $a \times b \in \mathbb{R}$.

M2 Associativity For all $a, b, c \in \mathbb{R}$,

$$a \times (b \times c) = (a \times b) \times c.$$

M3 Multiplicative identity For all $a \in \mathbb{R}$,

$$a \times 1 = a = 1 \times a.$$

M4 Multiplicative inverses For each $a \in \mathbb{R}^*$, there is a number $a^{-1} \in \mathbb{R}$ such that

$$a \times a^{-1} = 1 = a^{-1} \times a.$$

M5 Commutativity For all $a, b \in \mathbb{R}$,

$$a \times b = b \times a.$$

Property combining addition and multiplication in $\mathbb{R}$

D1 Distributivity For all $a, b, c \in \mathbb{R}$,

$$a \times (b + c) = (a \times b) + (a \times c).$$

The **additive identity** and **multiplicative identity** of $\mathbb{R}$ are 0 and 1, respectively.

The **additive inverse** or **negative** of a real number a is $-a$, and the **multiplicative inverse** or **reciprocal** of a non-zero real number a is a^{-1}.

5. A **field** is a set of numbers with addition and multiplication defined in such a way that they satisfy properties A1–A5, M1–M5 and D1, together with the property that the additive and multiplicative identities are different numbers.

The sets $\mathbb{R}$ and $\mathbb{Q}$ with ordinary addition and multiplication are fields, but $\mathbb{Z}$ is not.

6. A **polynomial** in x of **degree** n is an expression of the form

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

where $a_0, a_1, \ldots, a_n$ are numbers, called the **coefficients** of the polynomial, with $a_n \neq 0$.

A **real polynomial** is a polynomial with coefficients in $\mathbb{R}$.

7. A **polynomial equation** in x of degree n is an equation of the form $p(x) = 0$, where $p(x)$ is a polynomial in x of degree n.

Linear, **quadratic** and **cubic** equations (and polynomials) are polynomial equations (and polynomials) of degrees 1, 2 and 3, respectively.

8. The **roots** (or **zeros**) of a polynomial $p(x)$ are the solutions of the equation $p(x) = 0$.

A real polynomial of degree n has at most n distinct roots (some of which may be complex numbers).

9. **Factors of polynomials**

If a polynomial $p(x)$ can be expressed in the form

$$p(x) = s(x)t(x),$$

where $s(x)$ and $t(x)$ are polynomials whose degree is less than that of $p(x)$, then $s(x)$ and $t(x)$ are **factors** of $p(x)$.

Theorem A2 Factor Theorem (in $\mathbb{R}$)

Let $p(x)$ be a real polynomial, and let $\alpha \in \mathbb{R}$. Then $p(\alpha) = 0$ if and only if $x - \alpha$ is a factor of $p(x)$.

Theorem A3

Let $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ be a real polynomial, and suppose that $p(x)$ has n distinct real roots, $\alpha_1, \alpha_2, \ldots, \alpha_n$. Then

$$p(x) = a_n(x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n).$$

10. Suppose that

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = (x - \alpha_1)(x - \alpha_2)\cdots(x - \alpha_n),$$

where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are real numbers. Then

- $a_0 = (-1)^n \alpha_1 \alpha_2 \cdots \alpha_n$
- $a_{n-1} = -(\alpha_1 + \alpha_2 + \cdots + \alpha_n)$.

2 Complex numbers

11. A **complex number** is an expression of the form $x + iy$, where x and y are real numbers and $i^2 = -1$. The set of all complex numbers is denoted by $\mathbb{C}$.

The **real part** and **imaginary part** of a complex number $z = x + iy$ are the real numbers x and y, respectively. We write

$$\operatorname{Re} z = x \quad \text{and} \quad \operatorname{Im} z = y.$$

Two complex numbers are **equal** if their real parts are equal and their imaginary parts are equal.

12. The **Cartesian form** of a complex number z is the form $x + iy$, where $x, y \in \mathbb{R}$.

Any real number x can be written in the form $x + i0$, and any complex number of the form $x + i0$ is usually written simply as x. In this sense, $\mathbb{R}$ is a subset of $\mathbb{C}$.

The complex number $0 + i0$ is written as 0.

An **imaginary number** is a complex number of the form $0 + iy$ (where $y \neq 0$).

13. **Square roots of a negative real number**
For a positive real number d, the square roots of $-d$ are $\pm i\sqrt{d}$.

14. In the **complex plane** each complex number $z = x + iy$ is represented by the point (x, y).

Such a representation is also called an **Argand diagram**.

The **real axis** is the horizontal axis representing real numbers, and the **imaginary axis** is the vertical axis representing numbers of the form iy.

15. **Arithmetic of complex numbers in Cartesian form**

Arithmetic operations on complex numbers are carried out as for real numbers, except that we replace i^2 by -1 wherever it occurs.

Let $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$ be any complex numbers. Then the following operations can be applied.

Addition $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$

Subtraction $z_1 - z_2 = (x_1 - x_2) + i(y_1 - y_2)$

Multiplication

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i(x_2 y_1 + x_1 y_2)$$

16. The **complex conjugate** $\overline{z}$ of the complex number $z = x + iy$ is the complex number $x - iy$.

Properties of complex conjugates

Let z_1, z_2 and z be any complex numbers. Then:

1. $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$
2. $\overline{z_1 z_2} = \overline{z_1} \times \overline{z_2}$
3. $z + \overline{z} = 2\,\mathrm{Re}\, z$
4. $z - \overline{z} = 2i\,\mathrm{Im}\, z$.

17. The **modulus** $|z|$ of a complex number $z = x + iy$ is the distance from the point z in the complex plane to the origin:

$$|z| = \sqrt{x^2 + y^2}.$$

Properties of the modulus

1. $|z| \geq 0$ for any $z \in \mathbb{C}$, with equality only when $z = 0$.
2. $|z_1 z_2| = |z_1|\,|z_2|$ for any $z_1, z_2 \in \mathbb{C}$.

Conjugate–modulus properties

1. $|\overline{z}| = |z|$ for all $z \in \mathbb{C}$.
2. $z\overline{z} = |z|^2$ for all $z \in \mathbb{C}$.

18. Distance formula for $\mathbb{C}$ The distance between the points z_1 and z_2 in the complex plane is $|z_1 - z_2|$.

19. Quotient in Cartesian form

To simplify a quotient of complex numbers, multiply the numerator and denominator by the complex conjugate of the denominator.

20. Arithmetic in $\mathbb{C}$

The set $\mathbb{C}$ of complex numbers satisfies the eleven properties previously given for $\mathbb{R}$, as follows.

Properties for addition in $\mathbb{C}$

A1 Closure For all $z_1, z_2 \in \mathbb{C}$, $z_1 + z_2 \in \mathbb{C}$.

A2 Associativity For all $z_1, z_2, z_3 \in \mathbb{C}$,

$$z_1 + (z_2 + z_3) = (z_1 + z_2) + z_3.$$

A3 Additive identity For all $z \in \mathbb{C}$,

$$z + 0 = z = 0 + z.$$

A4 Additive inverses For each $z \in \mathbb{C}$, there is a number $-z \in \mathbb{C}$ such that

$$z + (-z) = 0 = (-z) + z.$$

A5 Commutativity For all $z_1, z_2 \in \mathbb{C}$,

$$z_1 + z_2 = z_2 + z_1.$$

Properties for multiplication in $\mathbb{C}$

M1 Closure For all $z_1, z_2 \in \mathbb{C}$, $z_1 \times z_2 \in \mathbb{C}$.

M2 Associativity For all $z_1, z_2, z_3 \in \mathbb{C}$,

$$z_1 \times (z_2 \times z_3) = (z_1 \times z_2) \times z_3.$$

M3 Multiplicative identity For all $z \in \mathbb{C}$,

$$z \times 1 = z = 1 \times z.$$

M4 Multiplicative inverses For each $z \in \mathbb{C} - \{0\}$, there is a number $z^{-1} \in \mathbb{C}$ such that

$$z \times z^{-1} = 1 = z^{-1} \times z.$$

M5 Commutativity For all $z_1, z_2 \in \mathbb{C}$,

$$z_1 \times z_2 = z_2 \times z_1.$$

Property combining addition and multiplication in $\mathbb{C}$

D1 Distributivity For all $z_1, z_2, z_3 \in \mathbb{C}$,

$$z_1 \times (z_2 + z_3) = (z_1 \times z_2) + (z_1 \times z_3).$$

The **additive identity** and **multiplicative identity** of $\mathbb{C}$ are $0 = 0 + 0i$ and $1 = 1 + 0i$, respectively.

The **additive inverse** (or negative) of $z = x + iy$ is $-z = -x - iy$, and the **multiplicative inverse** (or reciprocal) of $z = x + iy$ is

$$\frac{1}{z} = \frac{\overline{z}}{|z|^2} = \frac{x - iy}{x^2 + y^2}, \quad \text{for } z \neq 0.$$

21. The set $\mathbb{C}$ with addition and multiplication is a **field**, like $\mathbb{R}$ and $\mathbb{Q}$.

22. The complex numbers are not ordered, unlike the real numbers.

23. Polar form

A non-zero complex number z is in **polar form** if it is expressed as

$$z = r(\cos\theta + i\sin\theta),$$

where $r = |z|$ and θ is any angle (measured in radians anticlockwise) between the positive direction of the real axis and the line joining the origin to z.

An **argument**, $\arg z$, of the complex number z is such an angle θ.

The **principal argument**, $\operatorname{Arg} z$, of z is the value of $\arg z$ that lies in the interval $(-\pi, \pi]$.

24. Converting from polar form

To convert a complex number from polar form to Cartesian form, use the equations

$$x = r\cos\theta, \quad y = r\sin\theta.$$

25. Converting to polar form

To convert a non-zero complex number z from Cartesian form $x + iy$ to polar form $r(\cos\theta + i\sin\theta)$, first find the modulus r, using

$$r = |z| = \sqrt{x^2 + y^2}.$$

Then find the principal argument θ.

If z is either real or imaginary, then it lies on one of the axes and has principal argument 0, $\pi/2$, π or $-\pi/2$.

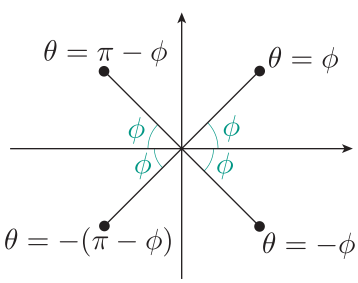
Otherwise, find the acute angle ϕ such that

$$\cos\phi = \frac{|x|}{r}.$$

This acute angle ϕ is the angle at the origin in the right-angled triangle formed by drawing the perpendicular from z to the real axis.

The principal argument θ can be found from the acute angle ϕ in either of the following ways:

- by sketching z in the complex plane, marking the acute angle ϕ on the sketch, and deducing the principal argument θ
- by using the appropriate formula (see below) to deduce θ from ϕ.



26. Product and quotient in polar form

- To multiply two (or more) complex numbers (all in polar form), multiply their moduli and add their arguments.
- To divide a complex number z_1 by a non-zero complex number z_2 (both in polar form), divide the modulus of z_1 by the modulus of z_2, and subtract the argument of z_2 from the argument of z_1.

To find the *principal* argument of a product or quotient, add or subtract integer multiples of 2π from the argument calculated, to obtain an angle in the interval $(-\pi, \pi]$.

> **Theorem A4 de Moivre's Theorem**
>
> If $z = \cos\theta + i\sin\theta$, then for any $n \in \mathbb{Z}$,
>
> $$z^n = (\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta.$$

27. The **roots** (or **zeros**) of a polynomial $p(z)$ are the solutions of the equation $p(z) = 0$.

28. Roots of a complex number

The ***n*th roots** of a complex number a are the solutions of the equation $z^n = a$.

Let $a = \rho(\cos\phi + i\sin\phi)$ be a complex number in polar form. Then, for any $n \in \mathbb{N}$, the equation $z^n = a$ has n solutions, given by

$$z = \rho^{1/n}\left(\cos\left(\frac{\phi}{n} + \frac{2k\pi}{n}\right) + i\sin\left(\frac{\phi}{n} + \frac{2k\pi}{n}\right)\right),$$

for $k = 0, 1, \dots, n-1$.

29. The ***n*th roots of unity** are the solutions of the equation $z^n = 1$. In the complex plane they are equally spaced around the unit circle.

30. Roots of complex polynomials

> **Theorem A5 The Fundamental Theorem of Algebra**
>
> Every polynomial equation
>
> $$a_n z^n + a_{n-1}z^{n-1} + \cdots + a_1 z + a_0 = 0,$$
>
> where $a_n, a_{n-1}, \dots, a_0 \in \mathbb{C}$ and $a_n \neq 0$, has at least one solution in $\mathbb{C}$.

The complex numbers form an **algebraically closed** system of numbers, since every polynomial equation with coefficients in this system has a solution in this system, which is not the case for the real numbers or the rational numbers.

31. **Factors of complex polynomials**

> **Theorem A6 Factor Theorem (in $\mathbb{C}$)**
>
> Let $p(z)$ be a polynomial with coefficients in $\mathbb{C}$, and let $\alpha \in \mathbb{C}$. Then $p(\alpha) = 0$ if and only if $z - \alpha$ is a factor of $p(z)$.
>
> **Theorem A7**
>
> Every polynomial
>
> $$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0,$$
>
> where $n \geq 1$ and the coefficients are in $\mathbb{C}$, with $a_n \neq 0$, has a factorisation
>
> $$p(z) = a_n (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_n),$$
>
> where the complex numbers $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the roots (not necessarily distinct) of $p(z)$.

32. **Complex roots of a polynomial with real coefficients**

> **Theorem A8**
>
> If $p(z)$ is a polynomial with *real* coefficients, then whenever α is a complex root of p, so is $\overline{\alpha}$.

A **complex conjugate pair** of factors α and $\overline{\alpha}$ combine to give a quadratic factor of $p(z)$ with real coefficients:

$$(z - \alpha)(z - \overline{\alpha}) = z^2 - (\alpha + \overline{\alpha})z + \alpha\overline{\alpha}.$$

33. **The complex exponential function**

If $z = x + iy$, then $e^z = e^x(\cos y + i \sin y)$.

> **Euler's Formula** $e^{iy} = \cos y + i \sin y$
>
> **Euler's Identity** $e^{i\pi} + 1 = 0$

34. **Exponential form**

A non-zero complex number z is in **exponential form** if it is expressed as

$$z = re^{i\theta}.$$

If $z = re^{i\theta}$, then $\overline{z} = re^{-i\theta}$.

35. In exponential form, **de Moivre's Theorem** (Theorem A4) becomes

$$(e^{i\theta})^n = e^{in\theta}, \quad \text{for all } \theta \in \mathbb{R} \text{ and all } n \in \mathbb{Z}.$$

3 Modular arithmetic

36. The Division Theorem describes the result of dividing an integer a by a positive integer n.

> **Theorem A9 Division Theorem**
>
> Let a and n be integers, with $n > 0$. Then there are unique integers q and r such that
>
> $$a = qn + r, \quad \text{with } 0 \leq r < n.$$

We say that dividing a by the **divisor** n gives the **quotient** q and **remainder** r.

37. **Congruence**

Let n be a positive integer. Two integers a and b are **congruent modulo n** if $a - b$ is a multiple of n, that is, if a and b have the same remainder on division by n.

We write this as the **congruence**

$$a \equiv b \pmod{n},$$

where n is the **modulus** of the congruence.

> **Theorem A10 Properties of congruences**
>
> Let n and m be positive integers, and let a, b, c, d be integers. The following properties hold.
>
> **Reflexivity** $a \equiv a \pmod{n}$
>
> **Symmetry** If $a \equiv b \pmod{n}$, then $b \equiv a \pmod{n}$.
>
> **Transitivity** If $a \equiv b \pmod{n}$ and $b \equiv c \pmod{n}$, then $a \equiv c \pmod{n}$.
>
> **Addition** If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $a + c \equiv b + d \pmod{n}$.
>
> **Multiplication** If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$, then $ac \equiv bd \pmod{n}$.
>
> **Powers** If $a \equiv b \pmod{n}$, then $a^m \equiv b^m \pmod{n}$.

38. Operations in $\mathbb{Z}_n$

For any integer $n \geq 2$,

$$\mathbb{Z}_n = \{0, 1, \ldots, n-1\}.$$

For a and b in $\mathbb{Z}_n$, the operations $+_n$ and $\times_n$ are:

- $a +_n b$ is the remainder of $a + b$ on division by n;
- $a \times_n b$ is the remainder of $a \times b$ on division by n.

The **modulus** for this arithmetic is the integer n.

Arithmetic modulo n is arithmetic carried out on the elements of the set $\mathbb{Z}_n$ using the operations $+_n$ and $\times_n$.

39. Arithmetic in $\mathbb{Z}_n$

Properties for addition in $\mathbb{Z}_n$ ($n \geq 2$)

A1 Closure For all $a, b \in \mathbb{Z}_n$, $a +_n b \in \mathbb{Z}_n$.

A2 Associativity For all $a, b, c \in \mathbb{Z}_n$,

$$(a +_n b) +_n c = a +_n (b +_n c).$$

A3 Additive identity For all $a \in \mathbb{Z}_n$,

$$a +_n 0 = a = 0 +_n a.$$

A4 Additive inverses For each $a \in \mathbb{Z}_n$, there is a number $b \in \mathbb{Z}_n$ such that

$$a +_n b = 0 = b +_n a.$$

A5 Commutativity For all $a, b \in \mathbb{Z}_n$,

$$a +_n b = b +_n a.$$

Properties for multiplication in $\mathbb{Z}_n$ ($n \geq 2$)

M1 Closure For all $a, b \in \mathbb{Z}_n$, $a \times_n b \in \mathbb{Z}_n$.

M2 Associativity For all $a, b, c \in \mathbb{Z}_n$,

$$(a \times_n b) \times_n c = a \times_n (b \times_n c)$$

M3 Multiplicative identity For all $a \in \mathbb{Z}_n$,

$$a \times_n 1 = a = 1 \times_n a.$$

M5 Commutativity For all $a, b \in \mathbb{Z}_n$,

$$a \times_n b = b \times_n a.$$

Combining addition and multiplication

D1 Distributivity For all $a, b, c \in \mathbb{Z}_n$,

$$a \times_n (b +_n c) = (a \times_n b) +_n (a \times_n c).$$

***Warning*:** Property M4 (multiplicative inverses) is missing from this list.

The **additive identity** and **multiplicative identity** of $\mathbb{Z}_n$ are 0 and 1, respectively.

40. The element b is a **multiplicative inverse** of a in $\mathbb{Z}_n$ if $a, b \in \mathbb{Z}_n$ and $a \times_n b = b \times_n a = 1$.

If an element of $\mathbb{Z}_n$ has a multiplicative inverse, then it has only one.

When it exists, we denote the multiplicative inverse b of an element a of $\mathbb{Z}_n$ by a^{-1} and refer to it as *the* multiplicative inverse of a in $\mathbb{Z}_n$.

***Warning*:** For composite n, $\mathbb{Z}_n$ contains non-zero elements that do not have a multiplicative inverse.

41. A **prime number** (or **prime**) is an integer greater than 1 whose only positive factors are 1 and itself; the first few primes are 2, 3, 5, 7, 11, 13, 17, and 19.

A **composite number** is an integer greater than 1 that is not a prime number; the first few composite numbers are 4, 6, 8, 9, 10, 12, 14, 15.

42. A **common factor** c of two integers a and b is a natural number that divides both a and b.

Two integers a and b are **coprime** (or **relatively prime**) if their only common factor is 1.

The **highest common factor** (**HCF**) of two integers a and b is their largest common factor.

Theorem A11

Let n and a be positive integers, with a in $\mathbb{Z}_n$. Then a has a multiplicative inverse in $\mathbb{Z}_n$ if and only if a and n are coprime.

43. Euclid's Algorithm is a method for finding the highest common factor of two positive integers.

Given an element a of $\mathbb{Z}_n$, we can apply Euclid's Algorithm to determine whether or not a and n are coprime; if they are coprime, then we can use the equations that arise to work out the multiplicative inverse of a, using a method known as **backwards substitution**.

44. Multiplication in $\mathbb{Z}_p$, for p prime

Multiplicative inverses in $\mathbb{Z}_p$

Let p be a prime number. Then every non-zero element in $\mathbb{Z}_p$ has a multiplicative inverse in $\mathbb{Z}_p$.

For multiplication in $\mathbb{Z}_p$, where p is a prime, we can add property M4 to the list of properties of multiplication in $\mathbb{Z}_n$, as follows.

M4 Multiplicative inverses For each $a \in \mathbb{Z}_p - \{0\}$ where p is a prime number, there is a number $a^{-1} \in \mathbb{Z}_p$ such that

$$a \times_p a^{-1} = 1 = a^{-1} \times_p a.$$

45. For each prime p, the set $\mathbb{Z}_p$ with arithmetic modulo p is a **field**.

46. Linear equations in modular arithmetic

Consider the linear equation

$$a \times_n x = b, \quad \text{where } a, b \in \mathbb{Z}_n.$$

When a and n are coprime:

The equation has the *unique* solution

$$x = a^{-1} \times_n b \quad \text{in } \mathbb{Z}_n.$$

To solve the equation use the inverse of a if known, or try different values of x (if the modulus n is fairly small), or try first spotting any solution in $\mathbb{Z}$ of the congruence $a \times x \equiv b \pmod{n}$: then the element of $\mathbb{Z}_n$ congruent to x modulo n is the required solution.

When a and n are not necessarily coprime:

Let d be the highest common factor of a and n.

- If d is not a factor of b, then the equation has no solutions in $\mathbb{Z}_n$.
- If d is a factor of b, then the equation has d solutions in $\mathbb{Z}_n$, given by

$$x = c, \quad x = c + \frac{n}{d}, \quad x = c + \frac{2n}{d}, \quad \ldots, \quad x = c + \frac{(d-1)n}{d},$$

where c is the unique solution in $\mathbb{Z}_{n/d}$ of the simpler equation

$$\frac{a}{d} \times_{\frac{n}{d}} x = \frac{b}{d}.$$

(The simpler equation has a unique solution since a/d and n/d are coprime.)

Unit A3 Mathematical language and proof

1 Mathematical statements

1. Mathematical statements

A mathematical **statement** (sometimes called a **proposition**) is an assertion that is either true or false, though we may not know which.

A **variable proposition** is a statement that is either true or false depending on the value of one or more variables.

A **theorem** is a mathematical statement that is true. Theorems are sometimes called *results*.

A **proposition** is a 'less important' theorem, and a **lemma** is a theorem that is used in the proof of other theorems.

A **corollary** is a theorem that follows from another theorem by a short additional argument.

2. The **negation**, 'not P', of a statement is a related statement that is true when the original statement is false, and false when the original statement is true.

3. Conjunctions and disjunctions

The **conjunction** of two statements P and Q is obtained by inserting the word 'and' between P and Q; it is true if both of P and Q are true, and false if at least one of P or Q is false.

The **disjunction** of two statements P and Q is obtained by inserting the word 'or' between P and Q; it is true if at least one of P or Q is true, and false if both of P and Q are false.

The negation of

- 'P and Q' is 'not P or not Q'
- 'P or Q' is 'not P and not Q'.

4. Implications

An **implication** is a statement of the form

if P, then Q.

Its **hypothesis** is the statement P, and its **conclusion** is the statement Q.

5. Ways of writing 'if P, then Q'

P implies Q	Q follows from P
$P \implies Q$	$Q \impliedby P$
P is sufficient for Q	Q is necessary for P
P only if Q	Q if P
	Q whenever P
	Q provided that P

6. Negation, converse and contrapositive

The **negation** of an implication is a *conjunction*: the negation of 'if P, then Q' is 'P and not Q'.

The **converse** of the implication 'if P, then Q' is the *implication* 'if Q, then P'. Knowledge of whether an implication is true or false tells you nothing about whether its converse is true or false.

The **contrapositive** of the implication 'if P, then Q' is the *implication* 'if not Q, then not P'. The contrapositive is equivalent to the original implication.

7. Equivalences

An **equivalence** is a statement of the form

if P, then Q, and if Q, then P.

It asserts that the implication 'if P, then Q' and its converse are *both* true. It is usually expressed more concisely as

P if and only if Q.

If the statement 'P if and only if Q' is true, then P and Q are either both true or both false.

8. Ways of writing 'P if and only if Q'

P is equivalent to Q

$P \iff Q$

P is necessary and sufficient for Q

$P \implies Q$ and $Q \implies P$

$P \implies Q$ and (not $P \implies$ not Q)

$\underset{\text{'if'}}{P \impliedby Q}$ and $\underset{\text{'only if'}}{P \implies Q}$

9. Universal and existential statements

The **universal quantifier** expresses the phrase 'for all' or an equivalent phrase such as 'every', 'for each', 'any'. It is often denoted by the symbol $\forall$.

A **universal statement** asserts that a certain property is true for all the elements of a given set. Universal statements usually start with a universal quantifier.

The **existential quantifier** expresses the phrase 'there is' or an equivalent phrase such as 'there exists', 'some', 'at least one'. It is often denoted by the symbol $\exists$.

An **existential statement** asserts that a certain property holds for at least one element in a given set. Existential statements usually start with an existential quantifier.

- The negation of a universal statement is an existential statement.
- The negation of an existential statement is a universal statement.

10. By convention, we omit the universal quantifier and write $P(x) \implies Q(x)$ for universal statements of the form $\forall x, P(x) \implies Q(x)$.

11. Negations

Statement	Negation
P	not P
P and Q	not P or not Q
P or Q	not P and not Q
If P, then Q	P and not Q
For all x, P	There exists an x such that not P
There exists an x such that P	For all x, not P

2 Direct proof

12. A **proof** of a mathematical statement is a logical argument that establishes that the statement is true.

A **direct proof** is a method of proof that involves a sequence of logical steps leading from known facts and assumptions directly to the statement to be proved.

13. Proof by exhaustion is a method of proof that can be used when there are only a small number of possibilities to consider: it involves proving a statement by considering each possibility in turn.

14. **Proving implications**

In general, to prove that the implication $P \implies Q$ is true, we start out by assuming that P is true, and build up a sequence of statements $P, P_1, P_2, \ldots, Q$, each of which follows from one or more statements further back in the sequence or from previous mathematical knowledge. If this can be achieved, then we have a proof of the implication $P \implies Q$.

15. A statement Q that is not an implication can be proved in a similar way to an implication $P \implies Q$. In this case the initial statement P is a statement that we know to be true from our previous knowledge.

16. **Proving equivalences**

To prove that the equivalence $P \iff Q$ is true, we have to prove *both* of the implications $P \implies Q$ and $Q \implies P$.

It may be possible to present the proof of both implications at once, by writing $P \iff P_1 \iff P_2 \iff \cdots \iff P_n \iff Q$.

17. **Proving existential and universal statements**

- An existential statement can often be proved by finding an object that satisfies the property in the statement.
- A universal statement about an infinite set can be proved *only* by using a general argument.

18. A **counterexample** to the implication $P(x) \implies Q(x)$ is a value of x that makes the statement $P(x)$ true and the statement $Q(x)$ false. It shows that the implication is false.

To prove that an equivalence $P \iff Q$ is false, show that at least one of the implications $P \implies Q$ and $Q \implies P$ is false, by providing a counterexample. That is, find a case where one of P or Q is true and the other is false.

A universal statement can be proved to be false using a counterexample.

19. **Proof by induction**

Principle of Mathematical Induction

To prove that a statement $P(n)$ is true for $n = 1, 2, \ldots$:

1. show that $P(1)$ is true
2. show that the implication

$$P(k) \implies P(k+1)$$

is true for $k = 1, 2, \ldots$.

The Principle of Mathematical Induction can be adapted to prove that a statement $P(n)$ is true for all integers n greater than or equal to some given integer other than 1.

When writing a proof by induction you should clearly identify the statement to be proved, $P(n)$, and structure your proof as follows:

- prove that $P(1)$ holds (or $P(n_0)$ for some initial $n_0 \neq 1$)
- write down $P(k)$ and assume that it holds for a general $k \geq 1$ (or $k \geq n_0$ where appropriate)
- state that we need to deduce $P(k+1)$, and write down $P(k+1)$
- deduce $P(k+1)$ from $P(k)$
- conclude that $P(n)$ holds for all natural numbers n (or for all $n \geq n_0$ where appropriate).

3 Indirect proof

20. A **proof by contradiction** of a statement Q is obtained by *assuming* that Q is *false* and deducing a **contradiction** (a statement that is definitely false).

To prove an implication 'if P, then Q' by contradiction, begin by assuming the implication is false and deduce a contradiction; that is, assume that P is true and Q is false, and deduce a contradiction.

21. A **proof by contraposition** of an implication 'if P, then Q' is obtained by proving the contrapositive implication

if not Q, then not P.

The contrapositive is equivalent to the original implication, but it may be easier to prove in some cases.

Useful results from Sections 2 and 3

22. Geometric Series Identity

> **Theorem A12**
>
> Let $a, b \in \mathbb{R}$ and let n be a positive integer. Then
>
> $$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}).$$

The following identity is a special case of the Geometric Series Identity.

$$x^n - 1 = (x-1)(x^{n-1} + x^{n-2} + \cdots + x + 1)$$

23. Fundamental Theorem of Arithmetic

> **Theorem A13**
>
> Every integer greater than 1 can be written as a product of prime numbers. The factorisation is unique up to the order of the factors.

24. Some properties of numbers

For $n = 1, 2, \dots,$

$$1 + 3 + \cdots + (2n-1) = n^2,$$
$$1 + 2 + \cdots + n = \tfrac{1}{2}n(n+1),$$
$$1^2 + 2^2 + \cdots + n^2 = \tfrac{1}{6}n(n+1)(2n+1),$$
$$1^3 + 2^3 + \cdots + n^3 = \tfrac{1}{4}n^2(n+1)^2.$$

> **Theorem A14**
>
> There are infinitely many prime numbers.

If an integer $n \geq 2$ is not divisible by any of the primes less than or equal to $\sqrt{n}$, then n is a prime number.

4 Equivalence relations

25. We say that $\sim$ is a **relation** on a set X if, whenever $x, y \in X$, the statement $x \sim y$ is either true or false.

If $\sim$ is a relation on a set X and $x \sim y$ is false for a particular pair of elements x and y in X, then we write $x \not\sim y$.

26. An **equivalence relation** on a set X is a relation $\sim$ satisfying the following three properties.

E1 Reflexivity For all x in X,

$$x \sim x.$$

E2 Symmetry For all x, y in X,

$$\text{if } x \sim y, \text{ then } y \sim x.$$

E3 Transitivity For all x, y, z in X,

$$\text{if } x \sim y \text{ and } y \sim z, \text{ then } x \sim z.$$

27. A relation is **reflexive**, **symmetric** or **transitive** if it has property E1, E2 or E3, respectively.

Relations exist with every combination of the three properties E1, E2 and E3.

28. A **partition** of a set X is a collection of non-empty subsets of X such that

- every pair of subsets in the collection is disjoint
- the union of all the subsets in the collection is the whole set X,

that is, *every* element of X belongs to *exactly one* of the subsets in the collection.

We say that such a collection of subsets **partitions** the set.

29. Equivalence classes

Let $\sim$ be an equivalence relation on a set X, and let $x \in X$. Then the **equivalence class** of x, denoted by $[\![x]\!]$, is the set

$$[\![x]\!] = \{y \in X : x \sim y\}.$$

It is the set of all elements in X related to x.

30. Equivalence classes and partitions

Theorem A16

The equivalence classes of an equivalence relation on a set X form a partition of the set X.

Proposition A17

The equivalence classes of an equivalence relation on a set X have the following property: if x and y are elements of X, then their equivalence classes $[\![x]\!]$ and $[\![y]\!]$ are either equal sets or disjoint sets.

31. Representatives of equivalence classes

An equivalence class can be denoted by $[\![x]\!]$ where x is any one of its elements.

A **representative** of an equivalence class $[\![x]\!]$ is a particular element x of the class used to denote the class.

A **set of representatives** for an equivalence relation $\sim$ on a set X is a set of elements of X that contains exactly one element from each equivalence class.

32. Congruence modulo n is a familiar example of an equivalence relation.

Theorem A15

For any integer $n > 1$, congruence modulo n is an equivalence relation on $\mathbb{Z}$.

A set of representatives for congruence modulo n is

$$\mathbb{Z}_n = \{0, 1, \ldots, n-1\}.$$

33. Congruence modulo 2π is the equivalence relation $\sim$ defined on $\mathbb{R}$ by

$$x \sim y \quad \text{if } x - y = 2\pi k \text{ for some integer } k.$$

We usually write $x \equiv y \pmod{2\pi}$.

Arithmetic modulo 2π is used when calculating the principal argument of a complex number.

Unit A4 Real functions, graphs and conics

1 Real functions and their graphs

1. Convention for real functions

When a real function is specified *only by a rule*, it is understood that the domain of the function is the set of all real numbers for which the rule is applicable, and the codomain of the function is $\mathbb{R}$.

2. A real function f with domain A maps each real number x in A to a corresponding real number $f(x)$. A vertical line drawn through any number on the x-axis crosses the graph of f at most once.

3. Basic functions

Sketches of basic functions are included in the Quick reference section on page 129.

(a) A **constant function** is a function of the form $f(x) = b$, where $b \in \mathbb{R}$.

The graph of the constant function $f(x) = b$ is the horizontal line with y-intercept b.

(b) A **linear function** is a function of the form $f(x) = ax + b$, where $a, b \in \mathbb{R}$ and $a \neq 0$.

The graph of the linear function $f(x) = ax + b$ is the straight line with gradient a and y-intercept b.

(c) A **quadratic function** is a function of the form $f(x) = ax^2 + bx + c$, where $a, b, c \in \mathbb{R}$ and $a \neq 0$.

The **completed-square form** of the rule of the quadratic function $f(x) = ax^2 + bx + c$ is

$$f(x) = a\,(x - \alpha)^2 + \beta,$$

where

$$\alpha = -\frac{b}{2a}, \quad \beta = c - \frac{b^2}{4a}.$$

The graph of this function is a parabola with vertex (α, β) and axis $x = \alpha$.

(d) A **cubic function** is a function of the form $f(x) = ax^3 + bx^2 + cx + d$, where $a, b, c, d \in \mathbb{R}$ and $a \neq 0$.

The graph of a cubic function crosses the x-axis once or three times, or (more rarely) crosses it once and 'touches' it once.

(e) A **linear rational function** is a function of the form

$$f(x) = \frac{ax+b}{cx+d},$$

where $a, b, c, d \in \mathbb{R}$, $c \neq 0$, and a and b are not both 0.

The graph of a linear rational function $f(x) = (ax+b)/(cx+d)$ is a rectangular hyperbola with a horizontal asymptote $y = a/c$ and a vertical asymptote $x = -d/c$.

(f) The **reciprocal function** $f(x) = 1/x$ is a linear rational function whose graph has the x- and y-axes as asymptotes.

(g) An **exponential function** is a function of the form $f(x) = a^x$, where $a \in \mathbb{R}$ is positive.

The graph of an exponential function always lies entirely above the x-axis and passes through the point $(0, 1)$.

***The* exponential function**, or exp, is the function $f(x) = e^x$ (see also **15.** on page 28).

(h) The **modulus function** is given by

$$|x| = \begin{cases} x, & \text{if } x \geq 0, \\ -x, & \text{if } x < 0. \end{cases}$$

The graph of the modulus function has a corner at the origin.

(i) The **integer part function** is $f(x) = \lfloor x \rfloor$.

For each x, the **integer part** $\lfloor x \rfloor$ of x is obtained by rounding down to the nearest integer. The rounding is always *down*, no matter whether x is positive or negative.

The graph of $f(x) = \lfloor x \rfloor$ consists of horizontal line segments with jumps between them.

4. Trigonometric functions

Sketches of trigonometric functions are included in the Quick reference section on page 129.

The graphs of the **sine** and **cosine** functions (sin and cos) lie between $y = -1$ and $y = 1$. They are periodic with period 2π and have exactly the same shape; the cosine graph is obtained by shifting the sine graph to the left by the distance $\pi/2$.

The **tangent** function (tan) is given by

$$\tan x = \frac{\sin x}{\cos x}.$$

Its graph is periodic with period π and has a vertical asymptote at each odd multiple of $\pi/2$.

The **cosecant** and **secant** functions (cosec and sec) are given by

$$\operatorname{cosec} x = \frac{1}{\sin x}, \quad \sec x = \frac{1}{\cos x}.$$

Their graphs both have period 2π and have exactly the same shape; the secant graph is obtained by shifting the cosecant graph to the left by the distance $\pi/2$. The cosecant graph has a vertical asymptote at each multiple of π.

The **cotangent** function (cot) is given by

$$\cot x = \frac{\cos x}{\sin x}.$$

Its graph is periodic with period π and has a vertical asymptote at each multiple of π.

5. Translations and scalings of graphs

An (α, β)-**translation** of a graph translates it by α units to the right and β units upwards.

Applying an (α, β)-translation to the graph of $y = f(x)$ gives the graph of

$$y = f(x - \alpha) + \beta.$$

A (λ, μ)-**scaling** of a graph, where $\lambda, \mu \neq 0$, scales (stretches) it by the factor λ in the x-direction and the factor μ in the y-direction.

Applying a (λ, μ)-scaling to the graph of $y = f(x)$ gives the graph of

$$y = \mu f\left(\frac{x}{\lambda}\right).$$

2 Graph sketching

6. Determining features of a graph

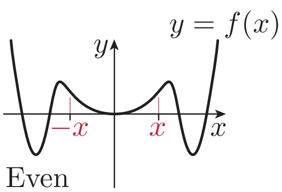
Domain:

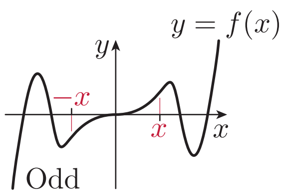
The domain of a real function specified just by a rule is the set of all real numbers, excluding any numbers that give an expression that is not defined.

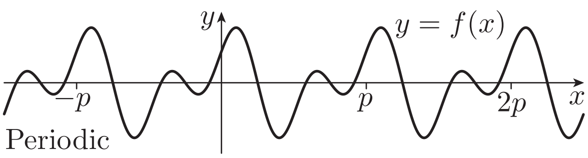
Symmetry features:

A function f is

- **even** if, for all x in the domain of f,
  $$f(-x) = f(x);$$
  its graph is unchanged when reflected in the y-axis
- **odd** if, for all x in the domain of f,
  $$f(-x) = -f(x);$$
  its graph is unchanged when rotated through the angle π about the origin
- **periodic** if there is a number p such that, for all x in the domain of f,
  $$f(x+p) = f(x);$$
  its graph is unchanged when translated along the x-axis by a distance p, but not when translated by any distance less than p.







Intercepts:

An **intercept** is a value of x or y at which the graph $y = f(x)$ of a function f meets the x- or y-axis, respectively.

The x-intercepts are the solutions (if there are any) of the equation $f(x) = 0$.

The y-intercept is the value $f(0)$, if this exists.

Intervals on which f is positive or negative:

A function f

- is **positive** on a particular interval I if $f(x) > 0$ for all $x \in I$
- is **negative** on a particular interval I if $f(x) < 0$ for all $x \in I$
- has a **zero** at x if $f(x) = 0$.

A **table of signs** for f is useful for finding the intervals on which f is positive or negative when f is a polynomial or a rational function.

Intervals on which f is increasing or decreasing:

A function f is

- **increasing** on an interval I, if for all $x_1, x_2 \in I$,
  $$\text{if } x_1 < x_2, \quad \text{then } f(x_1) \leq f(x_2)$$
- **strictly increasing** on an interval I, if for all $x_1, x_2 \in I$,
  $$\text{if } x_1 < x_2, \quad \text{then } f(x_1) < f(x_2)$$
- **decreasing** on an interval I, if for all $x_1, x_2 \in I$,
  $$\text{if } x_1 < x_2, \quad \text{then } f(x_1) \geq f(x_2)$$
- **strictly decreasing** on an interval I, if for all $x_1, x_2 \in I$,
  $$\text{if } x_1 < x_2, \quad \text{then } f(x_1) > f(x_2).$$

Increasing/decreasing criteria

- If $f'(x) > 0$ for all x in an interval I, then f is strictly increasing on I.
- If $f'(x) < 0$ for all x in an interval I, then f is strictly decreasing on I.

7. Stationary points

A **local maximum** is a point where the graph of a function changes from being strictly increasing to strictly decreasing.

A **local minimum** is a point where the graph of a function changes from being strictly decreasing to strictly increasing.

A **stationary point** of a differentiable function f is a value of x at which $f'(x) = 0$. The tangent to the graph of f at the point $(x, f(x))$ is horizontal.

A local maximum or minimum of a differentiable function is a stationary point.

First Derivative Test

Suppose that a is a stationary point of a differentiable function f, so that $f'(a) = 0$.

- If f' changes from positive to negative as x increases through a, then f has a **local maximum** at a.
- If f' changes from negative to positive as x increases through a, then f has a **local minimum** at a.
- If f' remains positive or remains negative as x increases through a (except at a itself, where $f'(a) = 0$), then f has a **horizontal point of inflection** at a.

A **table of signs** for f' is useful for finding the intervals on which f is increasing or decreasing; also for determining the nature of any stationary points for a polynomial or a rational function.

Second Derivative Test

Suppose that a is a stationary point of a differentiable function f, so that $f'(a) = 0$.

- If $f''(a) < 0$, then f has a local maximum at a.
- If $f''(a) > 0$, then f has a local minimum at a.

If $f''(a) = 0$, then the Second Derivative Test gives no result.

8. **Asymptotic behaviour** of a function f is the behaviour of the graph of $y = f(x)$ at the points of the graph for which the variable x or the variable y takes arbitrarily large values.

9. Asymptotes

An **asymptote** of a curve is a straight line that the curve approaches arbitrarily closely as the domain variable x or the codomain variable y (or both) take very large values.

A **vertical asymptote** is an asymptote with an equation of the form $x = a$.

A **horizontal asymptote** is an asymptote with an equation of the form $y = b$. The graph of a function may cross a *horizontal* asymptote.

For example, the curve below has a vertical asymptote such that

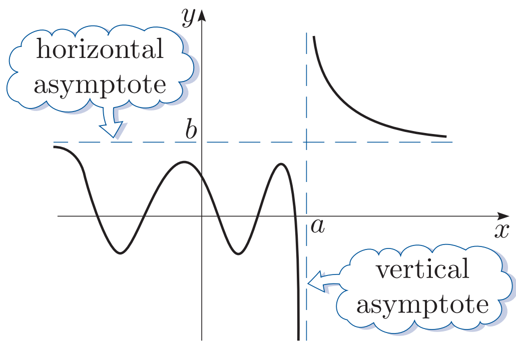
$$f(x) \to \infty, \quad \text{as } x \to a^+, \quad \text{and}$$
$$f(x) \to -\infty, \quad \text{as } x \to a^-.$$

Here a^+ and a^- indicate the direction of approach from values above and below a, respectively.

It also has a horizontal asymptote such that

$$f(x) \to b, \quad \text{as } x \to \infty, \quad \text{and}$$
$$f(x) \to b, \quad \text{as } x \to -\infty;$$

that is, $\quad f(x) \to b, \quad \text{as } x \to \pm\infty.$



10. Polynomial functions

A **polynomial function** of degree n is a function of the form

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

where $a_n \neq 0$.

Its **dominant term** is $a_n x^n$, the term with the highest power of x.

The coefficient of the dominant term is a_n.

A polynomial function f has no vertical asymptotes. Its asymptotic behaviour for large values of x is the same as that of its dominant term $a_n x^n$.

$a_n > 0$	$x \to \infty$	$x \to -\infty$
n even	$f(x) \to \infty$	$f(x) \to \infty$
n odd	$f(x) \to \infty$	$f(x) \to -\infty$

$a_n < 0$	$x \to \infty$	$x \to -\infty$
n even	$f(x) \to -\infty$	$f(x) \to -\infty$
n odd	$f(x) \to -\infty$	$f(x) \to \infty$

11. Rational functions

A **rational function** is a function of the form

$$f(x) = \frac{p(x)}{q(x)},$$

where p and q are polynomial functions.

It has vertical asymptotes at the values of x for which $q(x) = 0$ and $p(x) \neq 0$, if there are any such values.

It has at most one horizontal asymptote.

To identify any horizontal asymptote of a rational function, compare the dominant term of the numerator $p(x)$, say $a_n x^n$, with that of the denominator $q(x)$, say $b_m x^m$:

- if $n > m$, then the rational function has no horizontal asymptote
- if $n < m$, then the line $y = 0$ is a horizontal asymptote
- if $n = m$, then the line $y = c$ is a horizontal asymptote, where $c = a_n/b_m$.

12. Graph-sketching strategy

Strategy A3 Graph-sketching

To sketch the graph of a function f, determine the following features of f (where possible), and show these features in your sketch.

1. The domain of f.
2. Whether f is even, odd or periodic (or none of these).
3. The x-intercepts and the y-intercept of f, if any.
4. The intervals on which f is positive or negative.
5. The intervals on which f is increasing or decreasing, the nature of any stationary points, and the value of f at each of these points.
6. The asymptotic behaviour of f.

You may be able to obtain enough information without including all the steps.

3 New graphs from old

13. Extended graph-sketching strategy

To sketch the graph of a combination of two functions, one of which is a trigonometric function, it may help to exploit known features of the trigonometric functions and use other simple graphs as *construction lines* (which, if shown, must be dashed).

Strategy A4 Extended strategy

To sketch the graph of a function f, determine the features of f listed in steps 1–6 of Strategy A3 (where possible) in addition to step 7 below, and show these features in your sketch.

7. Any appropriate construction lines, and the points where f meets these lines.

14. A **hybrid function** has a rule that is defined by different formulas on different parts of its domain.

4 Hyperbolic functions

15. Properties of the exponential function

The function exp with rule $f(x) = e^x$ has the following properties:

- the domain of exp is $\mathbb{R}$
- exp is not even, odd or periodic
- $e^x > 0$ for all x in $\mathbb{R}$, so exp is positive on $\mathbb{R}$
- exp is its own derivative – that is, if $f(x) = e^x$, then $f'(x) = e^x$
- since $e^x > 0$ for all x in $\mathbb{R}$, exp is increasing on $\mathbb{R}$
- $e^0 = 1$, $e^x > 1$ for all $x > 0$ and $e^x < 1$ for all $x < 0$
- $e^{x+y} = e^x e^y$ for all x, y in $\mathbb{R}$
- $e^x \to \infty$ as $x \to \infty$ and $e^x \to 0$ as $x \to -\infty$
- if n is any positive integer, then $e^x/x^n \to \infty$ as $x \to \infty$, that is, e^x grows faster than any polynomial when x is sufficiently large.

16. Hyperbolic functions

Sketches of hyperbolic functions and a table of trigonometric and hyperbolic identities are included in the Quick reference section on pages 129 and 131.

- cosh is the **hyperbolic cosine function**, with rule
$$\cosh x = \frac{1}{2}\left(e^{x} + e^{-x}\right)$$
- sinh is the **hyperbolic sine function**, with rule
$$\sinh x = \frac{1}{2}\left(e^{x} - e^{-x}\right)$$
- tanh is the **hyperbolic tangent function**, with rule
$$\tanh x = \frac{\sinh x}{\cosh x}$$
- sech is the **hyperbolic secant function**, with rule
$$\operatorname{sech} x = \frac{1}{\cosh x}$$
- cosech is the **hyperbolic cosecant function**, with rule
$$\operatorname{cosech} x = \frac{1}{\sinh x}$$
- coth is the **hyperbolic cotangent function**, with rule
$$\coth x = \frac{1}{\tanh x}.$$

5 Conics

17. A **conic section**, or **conic**, is a curve obtained by slicing a double cone with a plane.

A **degenerate conic section**, or **degenerate conic**, is obtained when the slicing plane passes through the apex of the double cone. It may be a single point, a straight line or two intersecting straight lines.

A **non-degenerate conic section**, or **non-degenerate conic**, is a conic that is not degenerate. It may be a circle, an ellipse, a parabola or a hyperbola; sometimes a circle is considered to be a special type of ellipse.

18. Focus-directrix definitions of conics

The set of points P such that the distance of P from a fixed point is a constant multiple, e, of the distance of P from a fixed line is

- an **ellipse** if $0 < e < 1$
- a **parabola** if $e = 1$
- a **hyperbola** if $e > 1$.

The **focus** of the conic is the fixed point, the **directrix** is the fixed line, and the **eccentricity** is the constant multiple e.

19. Standard position

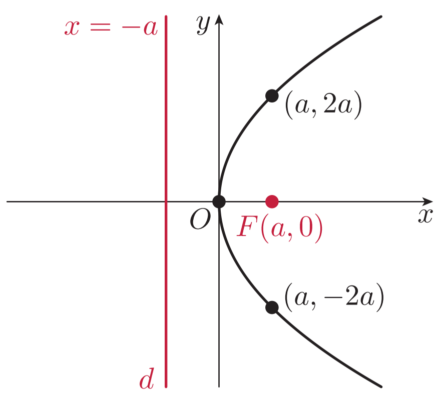
The non-degenerate conics in standard position are illustrated in the Quick reference section on page 135.

A (non-degenerate) conic is in **standard position** if it is positioned in the plane as follows.

- For a circle: its centre is at the origin.
- For an ellipse: its axes of symmetry are the x- and y-axes, and its largest width is along the x-axis.
- For a parabola: its axis of symmetry is the x-axis, it passes through the origin and its other points lie to the right of the origin.
- For a hyperbola: its axes of symmetry are the x- and y-axes, and it crosses the x-axis.

20. Parabola in standard position

A parabola in standard position has equation
$$y^2 = 4ax, \quad \text{where } a > 0.$$
It can be described by the parametric equations
$$x = at^2, \quad y = 2at \quad (t \in \mathbb{R}).$$
It has focus $(a, 0)$ and directrix $x = -a$; its axis is the x-axis and its vertex is the origin.



21. Ellipse in standard position

An ellipse in standard position has equation

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1,$$

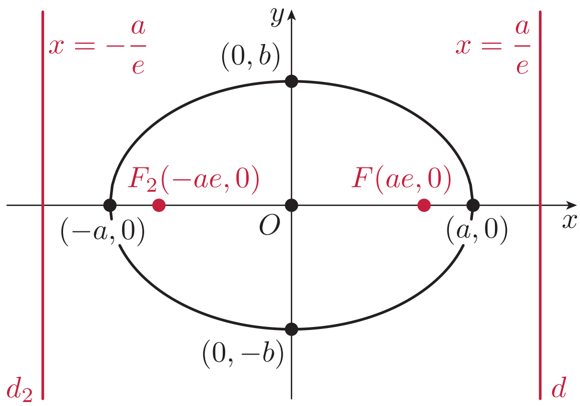
where $a \geq b > 0$, $b^2 = a^2(1 - e^2)$ and $0 < e < 1$; thus,

$$e = \sqrt{1 - \frac{b^2}{a^2}}.$$

It can be described by the parametric equations

$$x = a\cos t, \quad y = b\sin t \quad (t \in [0, 2\pi]).$$

It has foci $(\pm ae, 0)$ and directrices $x = \pm a/e$; its major axis is the line segment joining the points $(\pm a, 0)$, and its minor axis is the line segment joining the points $(0, \pm b)$.



A circle has no focus-directrix property, though it is sometimes considered to be an ellipse with eccentricity $e = 0$, a single focus and a directrix 'at infinity'.

22. Hyperbola in standard position

A hyperbola in standard position has equation

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1,$$

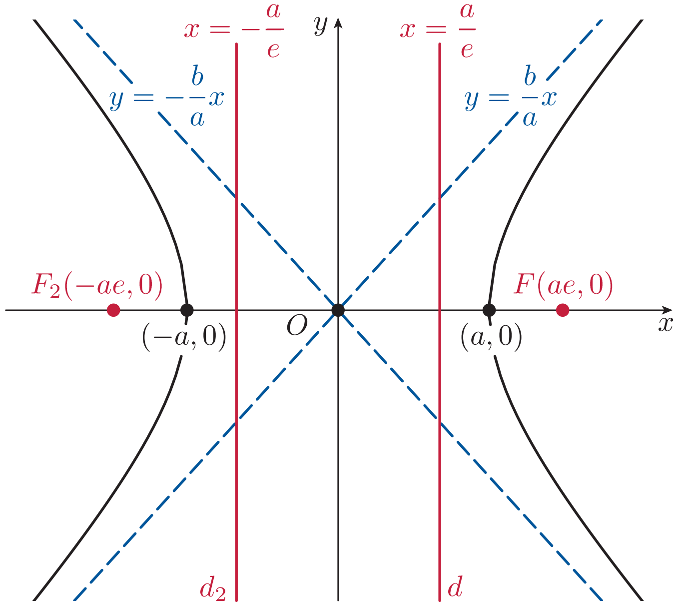
where $b^2 = a^2(e^2 - 1)$ and $e > 1$; thus

$$e = \sqrt{1 + \frac{b^2}{a^2}}.$$

It can be described by the parametric equations

$$x = a\sec t, \quad y = b\tan t$$
$$(t \in [-\pi, \pi] \text{ excluding } -\tfrac{\pi}{2} \text{ and } \tfrac{\pi}{2}).$$

It has foci $(\pm ae, 0)$ and directrices $x = \pm a/e$; it intersects the x-axis at the points $(\pm a, 0)$ and has two asymptotes $y = \pm(b/a)x$.



23. A **rectangular hyperbola** is a hyperbola whose asymptotes are perpendicular.

A rectangular hyperbola in standard position has $a = b$ and asymptotes $y = \pm x$.

24. General equation of a conic

When the empty set is defined to be a degenerate conic we have the following result.

Theorem A18

Any conic has an equation of the form

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0,$$

where A, B, C, F, G and H are real numbers, and A, B and C are not all zero.

Conversely, the set of all points in $\mathbb{R}^2$ whose coordinates (x, y) satisfy an equation of this form is a conic.

25. A **parametrisation** of a curve is a function of the form

$$\begin{aligned} f : I &\longrightarrow \mathbb{R}^2 \\ t &\longmapsto (f_1(t), f_2(t)) \end{aligned}$$

where I is an interval, and the image set of f comprises the points of the curve.

The **parameter** is the variable t, and f_1 and f_2 are real functions of the parameter t, both with domain I.

We write

$$f(t) = (f_1(t), f_2(t)), \quad t \in I.$$

The curve has **parametric equations**

$$x = f_1(t), \quad y = f_2(t).$$

A parametrisation of a given curve need not be unique. Different parametrisations of a curve may correspond to different modes of traversing the curve.

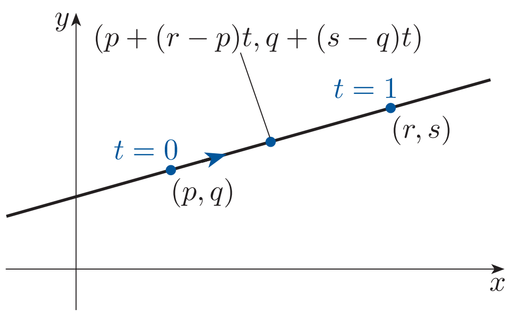
26. Parametrisation of a line

A parametrisation of the vertical line through the point $(p, 0)$ is

$$\alpha(t) = (p, t), \quad t \in \mathbb{R}.$$

A parametrisation of the line through the points (p, q) and (r, s) is

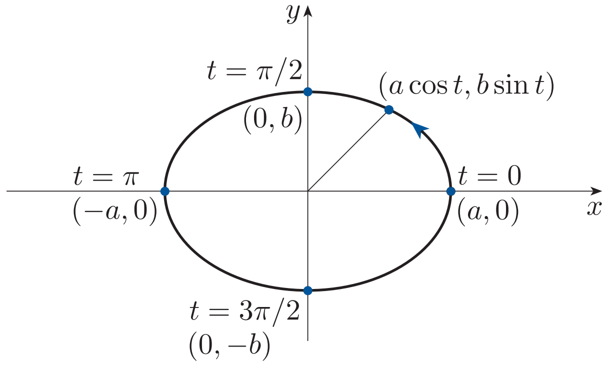
$$\alpha(t) = (p + (r - p)t, q + (s - q)t), \quad t \in \mathbb{R}.$$



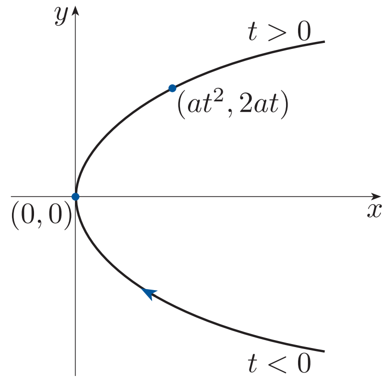
27. Parametrisations of conics in standard position

Circle: $\alpha(t) = (a\cos t, a\sin t), \quad t \in [0, 2\pi]$.

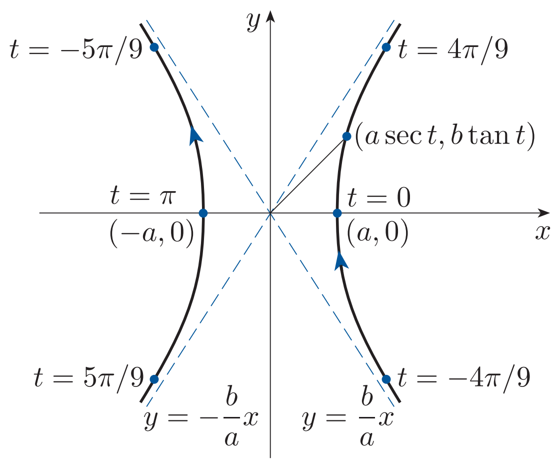
Ellipse: $\alpha(t) = (a\cos t, b\sin t), \quad t \in [0, 2\pi]$.



Parabola: $\alpha(t) = (at^2, 2at), \quad t \in \mathbb{R}$.



Hyperbola: $\alpha(t) = (a\sec t, b\tan t), \quad t \in [-\pi, \pi]$, excluding $-\pi/2$ and $\pi/2$.



Book B Group theory 1

Unit B1 Symmetry and groups

1 Symmetry in $\mathbb{R}^2$

1. A **plane figure** is any subset of the plane $\mathbb{R}^2$. A **bounded** plane figure is one that can be surrounded by a circle (of finite radius).

2. An **isometry** of the plane is a function $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ that preserves distances; that is, for all points $X, Y \in \mathbb{R}^2$, the distance between $f(X)$ and $f(Y)$ is the same as the distance between X and Y.

A **symmetry** of a plane figure F is an isometry that maps F to itself, that is, an isometry $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ such that $f(F) = F$.

3. An isometry of the plane is one of four types:

- a **rotation** rotates each point of the plane through the same angle about a particular point
- a **reflection** reflects each point of the plane in a particular line
- a **translation** moves each point of the plane by the same distance in the same direction
- a **glide-reflection** is a reflection in a line followed by a translation parallel to that line.

4. A symmetry of a bounded figure is one of the following.

- The **identity symmetry** (or **identity**), usually denoted by e: equivalent to doing nothing to a figure.
- A **rotation**: specified by a *centre* and an *angle of rotation*.
- A **reflection**: specified by a line – an *axis of symmetry*.

The identity symmetry can be regarded as a zero rotation or a zero translation.

5. The **trivial rotation** is the rotation through 0 radians; it is equal to the identity symmetry.

A **non-trivial rotation** is any rotation that is not equal to the trivial rotation.

When specifying a rotational symmetry, we measure angles anticlockwise (unless otherwise stated), and interpret negative angles as clockwise.

All the rotational symmetries of a bounded plane figure have the same centre of rotation (except that the identity symmetry can be regarded as a rotation about any point), and all the axes of symmetry of the figure pass through this centre.

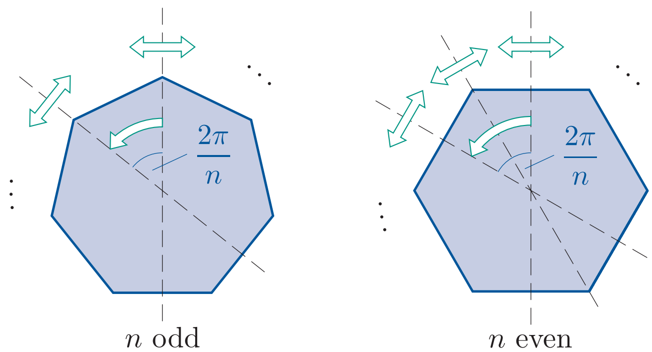
6. A **polygon** is a bounded plane figure with straight edges. A **regular polygon** is a polygon all of whose edges have the same length and all of whose angles are equal. A **regular n-gon** is a regular polygon with n edges.

7. Symmetries of a regular polygon

A regular n-gon has $2n$ symmetries: n rotations (through multiples of $2\pi/n$) and n reflections.

For odd values of n, each of the n axes of symmetry passes through a vertex and the midpoint of the opposite edge.

For even values of n, there are $n/2$ axes of symmetry that pass through opposite vertices and $n/2$ axes of symmetry that pass through the midpoints of opposite edges.



8. For any figure F, we denote the set of all symmetries of F by $S(F)$.

Every figure F has at least one symmetry, namely the identity symmetry. So, for every figure F, the set $S(F)$ of symmetries of F is non-empty.

9. Composition of symmetries

Order of composition is important. In general, if F is a figure and $f, g \in S(F)$, then $g \circ f$ may or may not be equal to $f \circ g$. That is, in general, composition of symmetries is not *commutative*.

Composition of symmetries of a bounded plane figure follows a standard pattern:

$\circ$	rotation	reflection
rotation	rotation	reflection
reflection	reflection	rotation

Composing a reflection with itself gives the identity symmetry e.

10. Properties of symmetries

Propositions B1, B2, B3 and B4

B1 Closure property for symmetries

The set of symmetries $S(F)$ of a plane figure F is **closed** under composition of functions; that is, for all elements f and g of $S(F)$, the composite $g \circ f$ is an element of $S(F)$.

B2 Associativity property for symmetries

Composition of symmetries is **associative**; that is, if F is a plane figure, then for all $f, g, h \in S(F)$,

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

B3 Identity property for symmetries

The set $S(F)$ of symmetries of a plane figure F contains a special symmetry e (the **identity symmetry**) such that, for each symmetry f in $S(F)$,

$$f \circ e = f = e \circ f.$$

B4 Inverses property for symmetries

Each symmetry f in the set $S(F)$ of symmetries of a plane figure F has an **inverse** symmetry f^{-1} in $S(F)$, such that

$$f \circ f^{-1} = e = f^{-1} \circ f.$$

11. Symmetries of the disc

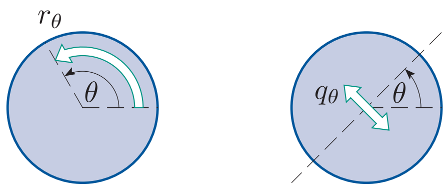
The set of symmetries of the disc is

$$S(\bigcirc) = \{r_\theta : \theta \in [0, 2\pi)\} \cup \{q_\theta : \theta \in [0, \pi)\},$$

where

r_θ is rotation through an angle θ about the centre (measured anticlockwise), for $\theta \in [0, 2\pi)$;

q_θ is reflection in the line through the centre at an angle θ to the horizontal (measured anticlockwise), for $\theta \in [0, \pi)$.



The zero rotation r_0 is the identity symmetry e. The symmetry q_0 is reflection in the horizontal axis and is not the identity symmetry.

12. Direct and indirect symmetries

The **direct** symmetries of a plane figure F are the symmetries of F that can be demonstrated with a paper model without lifting it out of the plane to turn it over. The set of direct symmetries of F is denoted by $S^+(F)$.

The **indirect** symmetries of F are the remaining symmetries: they are the symmetries that cannot be demonstrated with the paper model without lifting it out of the plane, turning it over and then replacing it in the plane.

For a *bounded* plane figure, the direct symmetries are rotations and the indirect symmetries are reflections.

Composition of direct and indirect symmetries follows a standard pattern:

$\circ$	direct	indirect
direct	direct	indirect
indirect	indirect	direct

Theorem B5

If a plane figure has a finite number of symmetries, then either

- all the symmetries are direct, or
- half of the symmetries are direct and half are indirect.

2 Representing symmetries

13. The **two-line symbol** representing a symmetry f of a polygon F that has vertices at locations labelled $1, 2, 3, \ldots, n$ is

$$\begin{pmatrix} 1 & 2 & 3 & \ldots & n \\ f(1) & f(2) & f(3) & \ldots & f(n) \end{pmatrix},$$

where $f(1), f(2), f(3), \ldots, f(n)$ are the labels of the locations to which f moves the vertices originally at the locations labelled $1, 2, 3, \ldots, n$, respectively.

We say that f is written in **two-line notation**.

The order of the columns in a two-line symbol is not important, though we usually use the natural order to aid recognition.

14. To find the *inverse* of a symmetry f, find the two-line symbol for f and turn it 'upside down', then rearrange the columns into the natural order.

3 Definition of a group

15. A **binary operation** on a set is a means of combining any two elements of the set.

Group axioms

Let G be a set and let $\circ$ be a binary operation defined on G. Then $(G, \circ)$ is a **group** if the following four **group axioms** hold.

G1 Closure For all g, h in G,

$$g \circ h \in G.$$

G2 Associativity For all g, h, k in G,

$$g \circ (h \circ k) = (g \circ h) \circ k.$$

G3 Identity There is an element e in G such that

$$g \circ e = g = e \circ g \quad \text{for all } g \text{ in } G.$$

(This element is an **identity** (**element**) for $\circ$ on G.)

G4 Inverses For each element g in G, there is an element h in G such that

$$g \circ h = e = h \circ g.$$

(The element h is an **inverse** (**element**) of g with respect to $\circ$.)

16. An **abelian** (or **commutative**) group $(G, \circ)$ has the additional property that

$$g \circ h = h \circ g \quad \text{for all } g, h \text{ in } G.$$

A **non-abelian** group is one that is not abelian.

We say elements g and h **commute** if $g \circ h = h \circ g$.

17. Order of a group

A group $(G, \circ)$ is **finite** if the set G is a finite set. A finite group $(G, \circ)$ is of **order** n if G has exactly n elements; we write $|G| = n$.

A group $(G, \circ)$ is **infinite** if the set G is an infinite set. It has **infinite order**.

18. Field definition in terms of groups

If F is a set, and $+$ and $\times$ are binary operations defined on F, then we say that $(F, +, \times)$ is a **field** if it has the following three properties.

- $(F, +)$ is an abelian group.
- $(F - \{0\}, \times)$ is an abelian group (where 0 is the identity element for $+$ on F).
- The distributive law

  $$x \times (y + z) = (x \times y) + (x \times z)$$

  holds for all $x, y, z \in F$.

19. Identity and inverses for sets of numbers

For a set of numbers, if the binary operation is:

- ordinary addition, then the only possible identity element is 0 and the only possible inverse of an element x is $-x$
- ordinary multiplication, then the only possible identity element is 1 and the only possible inverse of an element x is $1/x$.

The identity and inverses may be different for other binary operations such as those used in modular arithmetic.

20. You may assume without proof that the following binary operations are associative.

Standard associative binary operations

- Addition
- Multiplication
- Modular addition
- Modular multiplication
- Matrix addition
- Matrix multiplication
- Function composition

21. To show that a set and binary operation *do not* form a group, show that *any one* of the four group axioms fails.

22. A **Cayley table** for a finite set G and binary operation $\circ$ records the composites of the elements in G.

The elements of G are listed in the same order along the top and down the side. For any two elements $x, y \in G$, the composite $x \circ y$ is in the cell in the row labelled x and the column labelled y.

$$\begin{array}{c|ccc}
\circ & \cdots & y & \cdots \\
\hline
\vdots & & \vdots & \\
x & \cdots & x \circ y & \cdots \\
\vdots & & \vdots &
\end{array}$$

Note that x is on the left both in the composite and in the border of the table.

The **borders** of a Cayley table are the lists of elements along the top and down the side of the table. Its **body** is the rest of the table.

The **main diagonal** (or **leading diagonal**) of a Cayley table is the diagonal that goes from the top left to the bottom right. It contains the results of composing each element with itself.

A **group table** is a Cayley table of a group.

23. An element is **self-inverse** if it is its own inverse.

24. Identity and inverses in a Cayley table

Propositions B6 and B7

Let G be a finite set and let $\circ$ be a binary operation on G.

- The element e of G is an identity element for $\circ$ on G if and only if the row and column labelled e in a Cayley table for $(G, \circ)$ both repeat the table borders.
- The element h of G is an inverse of the element g of G if and only if e appears in a Cayley table for $(G, \circ)$ in the position that is in the row labelled g and column labelled h, and also in the position that is in the row labelled h and column labelled g.

Identifying inverses from a Cayley table

In a Cayley table for a set G and binary operation $\circ$ on G with an identity element e:

- wherever e occurs on the main diagonal, the corresponding element in the table borders is self-inverse
- wherever e occurs symmetrically with respect to the main diagonal, the corresponding elements in the table borders are inverses of each other.

25. Checking the group axioms for a small finite set

Using a Cayley table to check the group axioms

Let G be a finite set and let $\circ$ be a binary operation defined on G. Then $(G, \circ)$ is a group if and only if the Cayley table for $(G, \circ)$ has the following properties.

G1 Closure The table contains only elements of the set G; that is, no new elements appear in the body of the table.

G2 Associativity The operation $\circ$ is associative.
(This property is not easy to check from a Cayley table.)

G3 Identity A row and a column labelled by the same element repeat the table borders. This element is an identity element, e say.

G4 Inverses Each row contains the identity element e, occurring either on the main diagonal or symmetrically with another occurrence of e, with respect to the main diagonal. (For each such occurrence of e, the corresponding elements in the table borders are inverses of each other.)

26. $\mathbb{R}^*$, $\mathbb{Q}^*$ and $\mathbb{C}^*$ denote the sets of all non-zero real, rational and complex numbers, respectively.

$\mathbb{R}^+$ and $\mathbb{Q}^+$ denote the sets of all positive real and rational numbers, respectively.

27. Standard groups of numbers

$$(\mathbb{Z},+),\quad (\mathbb{Q},+),\quad (\mathbb{R},+),\quad (\mathbb{C},+),$$
$$(\mathbb{Q}^*,\times),\quad (\mathbb{R}^*,\times),\quad (\mathbb{C}^*,\times).$$

28. Groups from modular arithmetic

$\mathbb{Z}_n^*$ denotes the set of all *non-zero* integers in $\mathbb{Z}_n$.

U_n denotes the set of all integers in $\mathbb{Z}_n$ that are *coprime* to n.

Theorems B8 and B9

For each integer $n \geq 2$:

- the set $\mathbb{Z}_n$ is a group under $+_n$
- the set U_n is a group under $\times_n$.

Corollary B10

For each prime p, the set $\mathbb{Z}_p^*$ is a group under $\times_p$.

These are not the only groups that come from modular arithmetic.

4 Deductions from the group axioms

29. Axiom G2 (associativity) tells us that

- we can write a composite of any finite number of group elements with no brackets, without any ambiguity
- we can evaluate a composite of group elements by bracketing it however we wish, provided that we do not change the order in which the elements appear.

30. Uniqueness of the identity and inverses

Propositions B11 and B12

In any group:

- the identity element is unique
- each element has a unique inverse.

Notation Let g be an element of a group $(G,\circ)$ with identity element e. Then we denote the inverse of g by g^{-1}. So

$$g \circ g^{-1} = e = g^{-1} \circ g.$$

31. Properties of inverse elements

Proposition B13

Let g be an element of a group $(G,\circ)$. Then the inverse of g^{-1} is g, that is,

$$(g^{-1})^{-1} = g.$$

Proposition B14

Let x and y be elements of a group $(G,\circ)$. Then

$$(x \circ y)^{-1} = y^{-1} \circ x^{-1}.$$

Proposition B14 extends to composites of more than two group elements. For example, if x, y and z are elements of a group $(G,\circ)$, then

$$(x \circ y \circ z)^{-1} = z^{-1} \circ y^{-1} \circ x^{-1}.$$

***Warning*:** Note the reversed order.

32. Cancellation Laws

Proposition B15

In any group $(G,\circ)$ with elements a, b and x:

Left Cancellation Law

$$\text{if } x \circ a = x \circ b, \quad \text{then } a = b$$

Right Cancellation Law

$$\text{if } a \circ x = b \circ x, \quad \text{then } a = b.$$

33. Properties of group tables

Propositions B16, B17, B18 and B19

In a group table:

B16 the only elements in the body of the table are those that appear in the table borders

B17 the row and column corresponding to the identity element repeat the table borders

B18 each element of the group occurs exactly once in each row and exactly once in each column

B19 the identity element e occurs symmetrically with respect to the main diagonal.

Proposition B20

A group is abelian if and only if its group table is symmetrical with respect to the main diagonal.

5 Symmetry in $\mathbb{R}^3$

34. A **figure** in $\mathbb{R}^3$ is any subset of $\mathbb{R}^3$.
A **bounded** figure in $\mathbb{R}^3$ is one that can be surrounded by a sphere (of finite radius).

A **solid figure** (or **solid**) is a figure in $\mathbb{R}^3$ with non-zero height, non-zero width and non-zero depth.

35. Polyhedra

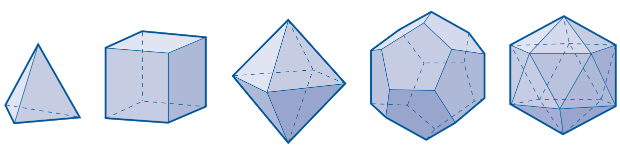
A **polyhedron** is a bounded solid whose faces are polygons.

A **convex** polyhedron is a polyhedron without dents or dimples or spikes: the line segment joining any two points that lie on different faces always lies inside the polyhedron.

A **regular polyhedron** (or **Platonic solid**) is a convex polyhedron in which all the faces are congruent regular polygons and at each vertex the same number of faces meet, arranged in the same way.

There are precisely five regular polyhedra:

- the **tetrahedron**, with four triangular faces
- the **cube**, with six square faces
- the **octahedron**, with eight triangular faces
- the **dodecahedron**, with twelve pentagonal faces
- the **icosahedron**, with twenty triangular faces.



36. An **isometry** of $\mathbb{R}^3$ is a function $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ that preserves distances.

A **symmetry** of a figure F in $\mathbb{R}^3$ is an isometry that maps F to itself; that is, an isometry $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ such that $f(F) = F$.

37. A symmetry of a bounded figure in $\mathbb{R}^3$ is one of the following.

- The **identity symmetry** e: equivalent to doing nothing to a figure.
- A **rotation**: specified by an axis of symmetry, together with a direction of rotation and an angle of rotation.
- A **reflection**: specified by the plane in which the reflection takes place.
- A **composite** of isometries of the types above – which may itself be of one of the types above, or (unlike with symmetries of plane figures) may not.

38. Symmetry group of a figure

Theorem B21

If F is a figure (in $\mathbb{R}^2$ or $\mathbb{R}^3$), then $S(F)$ is a group under function composition.

This group is called the **symmetry group** of F.

39. Direct and indirect symmetries

The **direct** symmetries of a figure in $\mathbb{R}^3$ are those that can be demonstrated with a model (that is, rotations).

The **indirect** symmetries are those that cannot be demonstrated physically with the model (that is, reflections or composites of a reflection and a rotation).

Theorem B22

If a figure (in $\mathbb{R}^2$ or in $\mathbb{R}^3$) has a finite number of symmetries, then either

- all the symmetries are direct, or
- half of the symmetries are direct and half are indirect.

If there are indirect symmetries, then they can all be obtained by composing any single indirect symmetry with all of the direct symmetries.

40. **Counting symmetries**

Strategy B1 Regular polyhedra

To determine the number of symmetries of a regular polyhedron, do the following.

1. Count the number of faces.
2. Count the number of symmetries of a face.
3. Then

$$\begin{pmatrix}\text{number of}\\ \text{symmetries of the}\\ \text{regular polyhedron}\end{pmatrix} = \begin{pmatrix}\text{number of}\\ \text{faces}\end{pmatrix} \times \begin{pmatrix}\text{number of}\\ \text{symmetries of a face}\end{pmatrix}.$$

Strategy B2 Non-regular polyhedra

To determine the number of symmetries of a non-regular polyhedron, do the following.

1. Select one type of face.

   (For two faces to be of the same type, it must be possible to place the polyhedron with either of the faces as its base and have it occupy the same space.)
2. Count the number of faces of this type.
3. Count the symmetries of a face of this type that give symmetries of the polyhedron.
4. Then

$$\begin{pmatrix}\text{number of}\\ \text{symmetries of}\\ \text{the polyhedron}\end{pmatrix} = \begin{pmatrix}\text{number of}\\ \text{faces of the}\\ \text{selected type}\end{pmatrix} \times \begin{pmatrix}\text{number of}\\ \text{symmetries of a face}\\ \text{of this type that}\\ \text{give symmetries of}\\ \text{the polyhedron}\end{pmatrix}.$$

Unit B2 Subgroups and isomorphisms

1 Subgroups

1. A **subgroup** of a group $(G, \circ)$ is a group $(H, \circ)$, where H is a subset of G.

A subgroup has the *same* binary operation as the original group.

2. The **trivial subgroup** of a group contains only one element: the identity element.

A **proper subgroup** of a group is any subgroup other than the whole group.

Every group $(G, \circ)$ with more than one element has at least two subgroups: the group $(G, \circ)$ itself, and the trivial subgroup $(\{e\}, \circ)$.

3. **Identities and inverses in subgroups**

Theorem B23

Let $(G, \circ)$ be a group with a subgroup $(H, \circ)$.

(a) The identity element of $(H, \circ)$ is the same as the identity element of $(G, \circ)$.

(b) For each element h of H, the inverse of h in $(H, \circ)$ is the same as its inverse in $(G, \circ)$.

4. **Subgroup properties**

Theorem B24 Subgroup test

Let $(G, \circ)$ be a group with identity element e, and let H be a subset of G. Then $(H, \circ)$ is a subgroup of $(G, \circ)$ if and only if the following three **subgroup properties** hold.

SG1 Closure For all x, y in H, the composite $x \circ y$ is in H.

SG2 Identity The identity element e of G is in H.

SG3 Inverses For each x in H, its inverse x^{-1} in G is in H.

To show that a subset H of G is *not* a subgroup of $(G, \circ)$, show that *any one* of the three subgroup properties fails.

5. Convention We often refer to a group $(G, \circ)$ just as the group G, provided this will not cause confusion. For instance, we often do this in the following situations:

- where there is an 'obvious' binary operation under which a set G is a group
- where the binary operation associated with a set G is clear from the context
- where we are discussing a general, abstract group and do not need to refer to the binary operation.

6. Subgroups of symmetry groups

> **Theorem B25**
>
> Let F be a figure in $\mathbb{R}^2$ or $\mathbb{R}^3$. Then the set $S^+(F)$ of direct symmetries of F is a subgroup of the symmetry group $S(F)$ of F.

> **Theorem B26**
>
> Let F be a figure in $\mathbb{R}^2$ or $\mathbb{R}^3$ and let A be a subset of F. Then the subset of $S(F)$ whose elements are all the symmetries of F that fix A is a subgroup of $S(F)$.

> **Strategy B3 Finding subgroups of symmetry groups**
>
> To find a subgroup of the symmetry group of a figure in $\mathbb{R}^2$ or $\mathbb{R}^3$, do *one* of the following.
>
> - Find the direct symmetries of the figure.
> - Modify the figure to restrict its symmetry; for example, introduce a pattern of lines or shapes. Then determine which of the symmetries of the original figure are symmetries of the new figure.
> - Find the symmetries of the figure that fix a particular vertex (or any other particular subset of the figure).

2 Order of a group element

7. Powers and multiples of a group element

Powers of an element x of a group $(G, \circ)$ are defined as follows. Let n be a positive integer. Then

$$x^0 = e, \quad \text{the identity element}$$
$$x^n = \underbrace{x \circ x \circ \cdots \circ x}_{n \text{ copies of } x}$$
$$x^{-n} = \underbrace{x^{-1} \circ x^{-1} \circ \cdots \circ x^{-1}}_{n \text{ copies of } x^{-1}}.$$

All powers of x are elements of G, since G is closed under $\circ$.

Multiples of an element x of a group $(G, +)$ are defined as follows. Let n be a positive integer. Then

$$0x = 0, \quad \text{the identity element}$$
$$nx = \underbrace{x + x + \cdots + x}_{n \text{ copies of } x}$$
$$-nx = \underbrace{(-x) + (-x) + \cdots + (-x)}_{n \text{ copies of } -x}.$$

All multiples of x are elements of G, since G is closed under $+$.

8. Multiplicative and additive notation

Feature	Multiplicative notation	Additive notation
Composite	$a \circ b$ or $a \times b$ or ab (or similar)	$a + b$ (or similar)
Identity	e or 1	0
Inverse	x^{-1}	$-x$
Power/multiple	x^n	nx

A **multiplicative** group is a group for which we use multiplicative notation.

An **additive** group is a group for which we use additive notation. Additive groups are always abelian, because addition is a commutative operation.

9. Index laws

Theorem B27 Multiplicative notation

Let x be an element of a group $(G, \circ)$, and let m and n be integers. The following hold.

(a) $x^m \circ x^n = x^{m+n}$

(b) $(x^m)^n = x^{mn}$

(c) $(x^n)^{-1} = x^{-n} = (x^{-1})^n$

Theorem B28 Additive notation

Let x be an element of a group $(G, +)$, and let m and n be integers. The following hold.

(a) $mx + nx = (m + n)x$

(b) $n(mx) = (nm)x$

(c) $-(nx) = (-n)x = n(-x)$

10. Order of a group element

Let x be an element of a group $(G, \circ)$.

The **order** of x is the *smallest* positive integer n such that $x^n = e$, if such an integer n exists. If it does exist, then we say that x has **finite order**.

If there is no positive integer n such that $x^n = e$, then x has **infinite order**.

Theorem B29

Let x be an element of a finite group $(G, \circ)$. Then x has finite order.

Theorem B30

If x is an element of a group $(G, \circ)$, then either x and x^{-1} have the *same* finite order, or they both have infinite order.

Theorem B31

Let x be an element of a group $(G, \circ)$.

(a) If x has finite order n, then the n powers

$$e, x, x^2, \ldots, x^{n-1}$$

are distinct, and these elements repeat indefinitely every n powers in the list of consecutive powers of x.

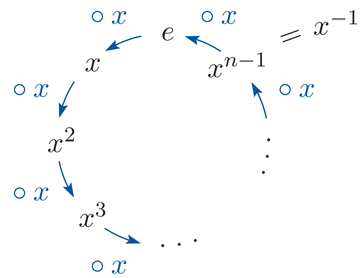
(b) If x has infinite order, then all the powers of x are distinct.

11. Order of the identity and order of self-inverse elements

Let $(G, \circ)$ be a group with identity element e.

- e has order 1.
- If the element x is self-inverse and $x \neq e$, then x has order 2.

12. The inverse of a group element x of finite order is the element that appears immediately before the identity element in the cycle of powers of x; that is, $x^{-1} = x^{n-1}$.



3 Cyclic subgroups and cyclic groups

13. Subgroup generated by an element

Let x be an element of a group $(G, \circ)$. The subset of G **generated** by x, denoted by $\langle x \rangle$, is the set of all powers of x:

$$\langle x \rangle = \{x^k : k \in \mathbb{Z}\}.$$

In additive notation, if x is an element of a group $(G, +)$, then the subset of G generated by x is the set of all multiples of x:

$$\langle x \rangle = \{kx : k \in \mathbb{Z}\}.$$

Theorem B32

Let x be an element of a group $(G, \circ)$. Then $(\langle x \rangle, \circ)$ is a subgroup of $(G, \circ)$.

This subgroup is called the **cyclic subgroup** of G **generated** by x.

The element x is a **generator** of $\langle x \rangle$.

Warning: $\langle x \rangle$ is a subgroup whereas x is an element.

14. Order and elements of a cyclic subgroup

Theorem B33

Let x be an element of a group.

(a) If x has finite order n, then the subgroup $\langle x \rangle$ has order n.

In multiplicative notation,

$$\langle x \rangle = \{e, x, x^2, \dots, x^{n-1}\}.$$

In additive notation,

$$\langle x \rangle = \{0, x, 2x, \dots, (n-1)x\}.$$

(b) If x has infinite order, then the subgroup $\langle x \rangle$ has infinite order.

In multiplicative notation,

$$\langle x \rangle = \{\dots, x^{-2}, x^{-1}, e, x, x^2, \dots\}.$$

In additive notation,

$$\langle x \rangle = \{\dots, -2x, -x, 0, x, 2x, \dots\}.$$

15. Some special cyclic subgroups

Let $(G, \circ)$ be a group with identity element e, and let $x \in G$.

- $\langle e \rangle = \{e\}$.
- If x is self-inverse and $x \neq e$, then $\langle x \rangle = \{e, x\}$.
- $\langle x^{-1} \rangle = \langle x \rangle$.

16. The set of **integer multiples** of any number x is

$$\begin{aligned} x\mathbb{Z} &= \{xk : k \in \mathbb{Z}\} \\ &= \{\dots, -2x, -x, 0, x, 2x, 3x, \dots\}. \end{aligned}$$

The group $(n\mathbb{Z}, +)$, for any integer n, is the subgroup of $(\mathbb{Z}, +)$ generated by n.

17. Cyclic groups

A group G is **cyclic** if there is an element $x \in G$ such that $G = \langle x \rangle$.

It is **non-cyclic** if there is no such element.

Theorem B34

Let G be a finite group of order n. Then G is cyclic if and only if G contains an element of order n.

Theorem B35

Every cyclic group is abelian.

***Warning*:** not all abelian groups are cyclic.

Theorem B36

Every subgroup of a cyclic group is cyclic.

18. The cyclic group $(\mathbb{Z}_n, +_n)$

Theorem B37

For each integer $n \geq 2$, the group $(\mathbb{Z}_n, +_n)$ is a cyclic group of order n. It is generated by the integer 1.

Theorem B38

Let m be a non-zero element of the group $(\mathbb{Z}_n, +_n)$. Then m has order n/d, where d is the HCF of m and n.

Corollary B39

Let m be a non-zero element of the group $(\mathbb{Z}_p, +_p)$, where p is prime. Then m has order p.

Corollary B40 Generators of $(\mathbb{Z}_n, +_n)$

Let $m \in \mathbb{Z}_n$. Then m is a generator of the group $(\mathbb{Z}_n, +_n)$ if and only if m is coprime to n.

Thus the generators of the group $(\mathbb{Z}_n, +_n)$ are the elements of the set U_n.

Theorem B41

The group $(\mathbb{Z}_n, +_n)$ has exactly one cyclic subgroup of order q for each positive factor q of n, and no other subgroups.

- The subgroup of order 1 is generated by 0.
- For each other factor q of n, the subgroup of order q is generated by d, where $qd = n$.

Remember that the factors of n include 1 and n.

4 Isomorphisms

19. **Isomorphic groups** are 'structurally identical' to each other.

Isomorphic groups

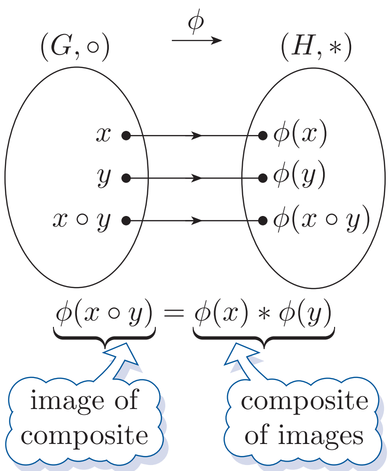
Two groups $(G, \circ)$ and $(H, *)$ are **isomorphic** if there exists a mapping $\phi : G \longrightarrow H$ with the following properties.

(a) ϕ is one-to-one and onto.

(b) For all $x, y \in G$,

$$\phi(x \circ y) = \phi(x) * \phi(y).$$

Such a mapping ϕ is called an **isomorphism**.



We use the symbol $\cong$ to denote the relation 'is isomorphic to'. If $(G, \circ)$ and $(H, *)$ are isomorphic, we write

$$(G, \circ) \cong (H, *)$$

or $G \cong H$ when the operations are clear.

Theorem B43

If two groups $(G, \circ)$ and $(H, *)$ are isomorphic, then either $(G, \circ)$ and $(H, *)$ are both finite, with the same order, or they are both infinite.

20. The **Klein four-group** V is a standard, abstract group isomorphic to $(S(\square), \circ)$.

***The* cyclic group of order n** is an abstract cyclic group of order n denoted by C_n.

21. **Isomorphism is an equivalence relation**

Theorem B44

The relation 'is isomorphic to' is an equivalence relation on the collection of all groups. That is, the following three properties hold.

Reflexivity Every group is isomorphic to itself.

Symmetry For any groups $(G, \circ)$ and $(H, *)$, if $(G, \circ)$ is isomorphic to $(H, *)$, then $(H, *)$ is isomorphic to $(G, \circ)$.

Transitivity For any groups $(G, \circ)$, $(H, *)$ and $(K, \vartriangle)$, if $(G, \circ)$ is isomorphic to $(H, *)$ and $(H, *)$ is isomorphic to $(K, \vartriangle)$, then $(G, \circ)$ is isomorphic to $(K, \vartriangle)$.

The collection of all groups can be partitioned into **isomorphism classes** such that two groups belong to the same isomorphism class if they are isomorphic, but belong to different classes otherwise.

22. **Showing that groups are isomorphic**

Strategy B4 Isomorphic groups

To show that two groups are isomorphic, try one of the following methods.

- Use facts that you know about the structures of the groups (such as whether they are cyclic, or abelian, and how many self-inverse elements they have), together with facts that you know about isomorphism classes.
- If the groups have small finite order, rearrange their Cayley tables to have the same pattern.
- If the groups are infinite or have large finite order, show algebraically that there is an isomorphism from one group to the other.

To help you identify a suitable rearrangement of a Cayley table, or an isomorphism, try using the properties in Theorems B45 and B46.

23. Properties of isomorphisms

Theorem B45

Let $(G, \circ)$ and $(H, *)$ be groups with identities e_G and e_H, respectively. Any isomorphism $\phi : (G, \circ) \longrightarrow (H, *)$ has the following properties.

(a) ϕ matches the identity elements:

$$\phi(e_G) = e_H.$$

(b) ϕ matches inverses: for each $g \in G$,

$$\phi(g^{-1}) = (\phi(g))^{-1}.$$

(c) ϕ matches powers of each element: for each $g \in G$ and each $k \in \mathbb{Z}$,

$$\phi(g^k) = (\phi(g))^k.$$

Theorem B46

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be an isomorphism. Let $g \in G$.

- If g has order n, then so does $\phi(g)$.
- If g has infinite order, then so does $\phi(g)$.

Theorem B47

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be an isomorphism. If K is a subgroup of $(G, \circ)$, then

$$\phi(K) = \{\phi(k) : k \in K\}$$

is a subgroup of $(H, *)$.

Theorem B48

Let $(G, \circ)$ and $(H, *)$ be isomorphic groups.

(a) If $(G, \circ)$ is abelian then so is $(H, *)$.

(b) If $(G, \circ)$ is cyclic then so is $(H, *)$.

24. Showing that groups are not isomorphic

Strategy B5 Non-isomorphic groups

To show that two groups are not isomorphic, try comparing their features. If any of the following are *not* the same for both groups, then they are not isomorphic:

- their orders
- whether they are abelian
- whether they are cyclic
- their numbers of self-inverse elements
- the numbers of different elements in the main diagonals of their Cayley tables
- their numbers of elements of a particular order.

***Warning*:** Even if all these are all the same, the two groups are not necessarily isomorphic.

25. Isomorphisms of cyclic groups

Theorems B49 and B50

Let $(G, \circ)$ and $(H, *)$ be cyclic groups generated by a and b, respectively. If $(G, \circ)$ and $(H, *)$ have the same finite order n, or they both have infinite order, then they are isomorphic, and an isomorphism is given by

$$\begin{aligned} \phi : G &\longrightarrow H \\ a^k &\longmapsto b^k, \end{aligned}$$

where $k = 0, 1, \ldots, n-1$ (for finite groups) or $k \in \mathbb{Z}$ (for infinite groups).

Strategy B6 Finding an isomorphism

To find an isomorphism between two finite cyclic groups $(G, \circ)$ and $(H, *)$ of the same order n, do the following.

1. Find a generator a of $(G, \circ)$ and a generator b of $(H, *)$.
2. Construct the isomorphism

$$\begin{aligned} \phi : G &\longrightarrow H \\ a^k &\longmapsto b^k \quad (k = 0, 1, \ldots, n-1). \end{aligned}$$

For additive cyclic groups the mapping $a^k \longmapsto b^k$ becomes $ka \longmapsto kb$.

Unit B3 Permutations

1 Permutations

1. A **permutation** of a finite set S is a one-to-one function from S to S.

We refer to the elements of S as the **symbols being permuted**. Usually we take $S = \{1, 2, \ldots, n\}$.

2. The **two-line form** of a permutation lists the symbols being permuted on one line, and the corresponding image of each symbol underneath it on a second line, enclosing the whole array in brackets.

For example, the permutation

$$f = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 6 & 8 & 3 & 1 & 2 & 7 & 5 \end{pmatrix}$$

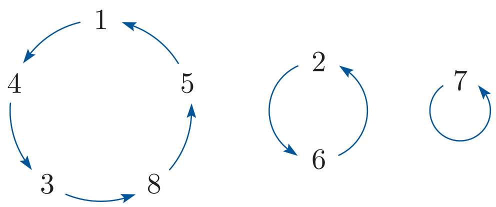
maps 1 to 4, maps 2 to 6, and so on.

3. The cycle form of a permutation

The permutation f above can also be written in *cycle form* as

$$f = (1\ 4\ 3\ 8\ 5)(2\ 6)(7).$$

This notation indicates that f maps the symbols as follows.



We say that f is the **product** of the the disjoint **cycles** $(1\ 4\ 3\ 8\ 5)$, $(2\ 6)$ and (7). Here **disjoint** means that each symbol belongs to only one cycle.

A permutation is written in **cycle form** when it is written as a product of disjoint cycles.

Theorem B51

Every permutation can be written in cycle form. The cycle form of a permutation is unique, apart from the choice of starting symbol in each cycle and the order in which the cycles are written.

Strategy B7 Finding the cycle form

To find the cycle form of a permutation f, do the following.

1. Choose any symbol (such as 1) and find its image under f, then find the image of its image under f, and so on, until you encounter the starting symbol again.
2. Write these symbols as a cycle.
3. Repeat the process starting with any symbol that has not yet been placed in a cycle, until all the symbols have been placed in cycles.

In step 3 it is sensible to start with the next available symbol in numerical order.

4. A symbol is **fixed** by a permutation if the permutation maps that symbol to itself; it forms a cycle containing just that symbol.

5. The **identity permutation** of a set S is the permutation of S that fixes every symbol.

6. Cycle form conventions

- We usually write the smallest symbol in each cycle first, and arrange the cycles with their smallest symbols in increasing order.
- When it is clear from the context which set of symbols is being permuted, we omit fixed symbols from the cycle form.
- When working with permutations in cycle form, we denote the identity permutation by e.

7. Composing permutations

A composite of permutations is a permutation.

Composition of permutations is not *commutative* in general: if f and g are permutations, then $g \circ f$ and $f \circ g$ may not be equal.

Strategy B8 Finding $g \circ f$

To find the composite $g \circ f$ of two permutations written in cycle form, do the following.

1. Start with the smallest symbol, 1 say. Find the symbol that is the image of 1 under f, then find the image of that symbol under g, and write the result, x say, next to 1 in a cycle:
   $$(1\ x\ \ldots.$$
2. Starting with the symbol x, repeat the process to obtain the next symbol in the cycle.
3. Continue repeating the process until the next symbol found is the original symbol 1. This completes the cycle.
4. Starting with the smallest symbol not yet placed in a cycle, repeat steps 1 to 3 until every symbol has been placed in a cycle.
5. Usually, rewrite the cycle form omitting the cycles containing a single symbol, if there are any.

A permutation is equal to the composite of its disjoint cycles; for example, if

$$f = (1\ 3)(2\ 4\ 9\ 6)(5\ 7\ 8),$$

then

$$f = (1\ 3) \circ (2\ 4\ 9\ 6) \circ (5\ 7\ 8).$$

8. Inverse of a permutation

The **inverse** f^{-1} of a permutation f undoes what f does; that is, if f maps x to y, then f^{-1} maps y to x.

Strategy B9 Finding f^{-1}

To find the inverse of a permutation written in cycle form, do the following.

1. Reverse the order of the symbols in each cycle.
2. Then, usually, rewrite each cycle so that the smallest symbol is first.

2 Permutation groups

9. The symmetric group S_n

Theorem B52

The set S_n of all permutations of the set $\{1, 2, 3, \ldots, n\}$ is a group under function composition.

This group S_n is called the **symmetric group of degree n**.

Theorem B53

The symmetric group S_n has order $n!$.

10. A **permutation group** (or **group of permutations**) is a subgroup of the group S_n, for some positive integer n.

11. Symmetry groups as subgroups of S_n

When the symmetries of a figure are represented by permutations they form a subgroup of a symmetric group.

If a figure has n vertices (or other features, such as edges or faces), and we label the locations of these vertices (or other features) with the symbols $1, 2, \ldots, n$, then the permutations of these symbols that represent the symmetries of the figure form a subgroup of the group S_n.

More generally, if we label the locations of the vertices (or other features) of a figure with some or all of the symbols from the set $\{1, 2, \ldots, n\}$, then the permutations of these symbols that represent the symmetries of the figure form a subgroup of the symmetric group S_n. Any symbols in $\{1, 2, \ldots, n\}$ that are not used to label the figure are taken to be fixed.

12. An **r-cycle** or a **cycle of length r** is a permutation whose cycle form consists of a single cycle permuting r symbols (with all other symbols, if there are any, fixed).

A **transposition** is a 2-cycle.

13. Two permutations in S_n have the **same cycle structure** if their cycle forms contain the same number of disjoint r-cycles for each natural number r.

14. Elements of S_3 and S_4

Cycle structure	Elements of S_3	Description
e	e	identity
$(-\ -)$	$(1\ 2)$, $(1\ 3)$, $(2\ 3)$	transpositions
$(-\ -\ -)$	$(1\ 2\ 3)$, $(1\ 3\ 2)$	3-cycles

Cycle structure	Elements of S_4	Description
e	e	identity
$(-\ -)$	$(1\ 2)$, $(1\ 3)$, $(1\ 4)$, $(2\ 3)$, $(2\ 4)$, $(3\ 4)$	transpositions
$(-\ -\ -)$	$(1\ 2\ 3)$, $(1\ 3\ 2)$, $(1\ 2\ 4)$, $(1\ 4\ 2)$, $(1\ 3\ 4)$, $(1\ 4\ 3)$, $(2\ 3\ 4)$, $(2\ 4\ 3)$	3-cycles
$(-\ -\ -\ -)$	$(1\ 2\ 3\ 4)$, $(1\ 4\ 3\ 2)$, $(1\ 2\ 4\ 3)$, $(1\ 3\ 4\ 2)$, $(1\ 3\ 2\ 4)$, $(1\ 4\ 2\ 3)$	4-cycles
$(-\ -)(-\ -)$	$(1\ 2)(3\ 4)$, $(1\ 3)(2\ 4)$, $(1\ 4)(2\ 3)$	two 2-cycles

15. Order of a permutation

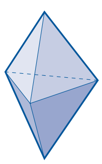
Theorem B54

An r-cycle has order r.

Theorem B55

The order of a permutation is the least common multiple of the lengths of its cycles.

16. The **double tetrahedron** is the solid formed by sticking together two regular tetrahedrons of the same size.



3 Even and odd permutations

17. Expressing a cycle as a composite of transpositions

Strategy B10 Transpositions

To express a cycle $(a_1\ a_2\ a_3\ \dots\ a_r)$ as a composite of transpositions, do the following.

Write the transpositions

$$(a_1\ a_2), (a_1\ a_3), (a_1\ a_4), \dots, (a_1\ a_r)$$

in reverse order and form their composite. That is,

$$(a_1\ a_2\ a_3 \dots\ a_r) = (a_1\ a_r) \circ (a_1\ a_{r-1}) \circ \dots \circ (a_1\ a_3) \circ (a_1\ a_2).$$

This strategy does not produce a *unique* expression for a cycle as a composite of transpositions. However, for an r-cycle it always produces a composite of $r - 1$ transpositions.

18. Expressing a permutation as a composite of transpositions

Theorem B56

Every permutation can be expressed as a composite of transpositions.

To express a permutation as a composite of transpositions, apply Strategy B10 to each of its cycles.

19. Parity of a permutation

Theorem B58 Parity Theorem

A permutation cannot be expressed both as a composite of an even number of transpositions and as a composite of an odd number of transpositions.

An **even** permutation can be expressed as a composite of an even number of transpositions.

An **odd** permutation can be expressed as a composite of an odd number of transpositions.

The **parity** of a permutation is its evenness/oddness.

20. Parity of a cycle

Theorem B59

An r-cycle is

- an even permutation, if r is odd
- an odd permutation, if r is even.

In particular, a transposition is an odd permutation and the identity permutation is an even permutation.

21. Finding the parity of a permutation

Strategy B11 Parity of a permutation

To determine the parity of a permutation, do the following.

1. Express the permutation as a composite of cycles (either disjoint or not).
2. Find the parity of each cycle, using the rule

$$\text{an } r\text{-cycle is } \begin{cases} \text{even}, & \text{if } r \text{ is odd}, \\ \text{odd}, & \text{if } r \text{ is even}. \end{cases}$$

3. Combine the even and odd parities using the following table.

+	even	odd
even	even	odd
odd	odd	even

22. Parity of an inverse permutation

Theorem B60

A permutation and its inverse have the same parity.

23. The alternating group A_n

Theorem B61

The set A_n of all even permutations of the set $\{1, 2, 3, \dots, n\}$ is a subgroup of the symmetric group S_n.

This group A_n is called the **alternating group of degree** $\boldsymbol{n}$.

Theorem B62

For each integer $n \geq 2$, the alternating group A_n has order $\frac{1}{2}(n!)$.

24. The **alternating group** $\boldsymbol{A_4}$ comprises the twelve elements

$$\begin{array}{cccc} e, & (1\ 2)(3\ 4), & (1\ 3)(2\ 4), & (1\ 4)(2\ 3), \\ (1\ 2\ 3), & (1\ 2\ 4), & (1\ 3\ 4), & (2\ 3\ 4), \\ (1\ 3\ 2), & (1\ 4\ 2), & (1\ 4\ 3), & (2\ 4\ 3). \end{array}$$

4 Conjugacy in S_n

25. Conjugate permutations

The permutation y is a **conjugate** in S_n of the permutation x if there is a permutation g in S_n such that

$$y = g \circ x \circ g^{-1}.$$

We say that:

- g **conjugates** x to y
- y is the **conjugate** of x by g
- g is a **conjugating permutation**
- x and y are **conjugates**, or **conjugate permutations**.

26. Renaming method

To find the conjugate $y = g \circ x \circ g^{-1}$, where x and g are permutations in S_n, replace each symbol in the cycle form of x by its image under g.

For example,

$$\begin{array}{rl} x = & (1\ 2\ 5)(3\ 4) \\ g \downarrow & \ \ \downarrow\ \downarrow\ \downarrow\ \ \ \downarrow\ \downarrow \\ y = & (3\ 2\ 4)(5\ 1) \end{array} \qquad \text{where } g = (1\ 3\ 5\ 4).$$

27. Finding conjugating permutations

Strategy B12 Finding a conjugating permutation

To find a permutation g such that $y = g \circ x \circ g^{-1}$, where x and y are permutations with the same cycle structure, do the following.

1. Match up the cycles of x and y (including 1-cycles) so that cycles of equal lengths correspond.

$$\begin{array}{rl} x = & (* \ * \ \cdots \ *)(* \ * \ \cdots \ *)\cdots(*)(*) \\ g\downarrow & \ \ \downarrow \downarrow \cdots \downarrow \ \ \downarrow \downarrow \cdots \downarrow \ \cdots \downarrow \ \downarrow \\ y = & (* \ * \ \cdots \ *)(* \ * \ \cdots \ *)\cdots(*)(*) \end{array}$$

2. Read off the two-line form of the renaming permutation g from this diagram. (Usually, write g in cycle form.)

For example, with $x = (1\ 2\ 4)(3\ 5)$ and $y = (1\ 4)(2\ 5\ 3)$ in S_5, we obtain

$$\begin{array}{rl} x = & (1\ 2\ 4)(3\ 5) \\ g\downarrow & \ \ \downarrow \downarrow \downarrow \ \ \downarrow \downarrow \\ y = & (2\ 5\ 3)(1\ 4), \end{array}$$

which gives $g = (1\ 2\ 5\ 4\ 3)$.

In general, to find further permutations g such that $y = g \circ x \circ g^{-1}$, change the starting symbols in the cycles of y, and rearrange the cycles of y, if possible. (For example, above there are five other ways to write y including $y = (5\ 3\ 2)(1\ 4)$ and $y = (5\ 3\ 2)(4\ 1)$.)

28. Conjugate permutations and cycle structure

Theorem B64

Two permutations in the symmetric group S_n are conjugate in S_n if and only if they have the same cycle structure.

29. Conjugate subgroups in S_n

Let H be a subgroup of S_n, and let $g \in S_n$. Then

$$g \circ H \circ g^{-1} = \{g \circ h \circ g^{-1} : h \in H\}.$$

That is, $g \circ H \circ g^{-1}$ is the set obtained by conjugating every element of H by g.

Theorem B65

Let H be a subgroup of S_n, and let $g \in S_n$. Then $g \circ H \circ g^{-1}$ is also a subgroup of S_n.

This subgroup is called the **conjugate subgroup** of H by g. We say that $g \circ H \circ g^{-1}$ is a **conjugate subgroup** of H in S_n.

30. Finding conjugate subgroups in S_n

Strategy B13 Renaming for a subgroup

To find the subgroup $g \circ H \circ g^{-1}$, given a subgroup H and an element g of S_n, do the following.

For each $h \in H$, find $g \circ h \circ g^{-1}$ by using g to 'rename' the symbols in h.

5 Subgroups of S_4

31. The number of subgroups of S_4 of each order is given below.

Order	Number of subgroups	Description
1	1	$\{e\}$
2	9	all cyclic
3	4	all cyclic
4	7	3 cyclic; 4 Klein
6	4	all isomorphic to $S(\triangle)$
8	3	all isomorphic to $S(\square)$
12	1	A_4
24	1	S_4

32. The symmetric group S_4 represents the symmetry group $S(\text{tet})$ of a regular tetrahedron with the vertices labelled 1, 2, 3 and 4.

The subgroup A_4 of S_4 represents the group $S^+(\text{tet})$ of direct symmetries of the tetrahedron.

33. Finding subgroups of groups in general

To find the cyclic subgroups of a group:
If the group is of reasonably small order, list its elements according to their orders and then systematically find the subgroups generated by these elements. (Any element of order m in a cyclic subgroup of order m will generate that same cyclic subgroup.)

To find a subgroup of a symmetric group S_n:

- Find a symmetry group whose elements can be represented by permutations in S_n.
- Find all permutations that fix one or more symbols.

To find subgroups of a symmetry group:
Use Strategy B3. (See **6.** on page 39.)

It is usually difficult to find *all* the non-cyclic subgroups of a symmetric group or a symmetry group using these methods.

6 Cayley's Theorem

34. Representing a finite group as a permutation group

Theorem B66

Let $(G, *)$ be a finite group. For each element x of G, let p_x be the permutation whose two-line form has as its first line the column headings of the group table of $(G, *)$ and as its second line the row labelled x in the group table. Let

$$P = \{p_x : x \in G\}.$$

Then $(P, \circ)$ is a permutation group isomorphic to $(G, *)$.

Theorem B67 Cayley's Theorem

Every finite group is isomorphic to a permutation group.

Unit B4 Lagrange's Theorem and small groups

1 Lagrange's Theorem

1. Lagrange's Theorem allows us to write down all the *possible* orders for subgroups of a finite group G – these are all the positive divisors of the order of G. Thus, if the number m does *not* divide the order of G, then G does not have a subgroup of order m.

Theorem B68 Lagrange's Theorem

Let G be a finite group and let H be any subgroup of G. Then the order of H divides the order of G.

***Warning*:** The converse of Lagrange's Theorem is *false*. Lagrange's Theorem does *not* assert that if m is a positive divisor of the order of a group G, then G has a subgroup of order m.

2. Orders of group elements

Corollary B69 to Lagrange's Theorem

Let g be an element of a finite group G. Then the order of g divides the order of G.

3. Groups of prime order

Corollary B70 to Lagrange's Theorem

Let G be a group of prime order. Then G is cyclic, and every element of G other than the identity element is a generator of G.

Corollary B71 to Lagrange's Theorem

If G is a group of prime order p, then G is isomorphic to the cyclic group $(\mathbb{Z}_p, +_p)$.

Corollary B72 to Lagrange's Theorem

If G is a group of prime order, then the only subgroups of G are the trivial subgroup and G itself.

2 Groups of small order

4. Notation convention for abstract groups

An **abstract group** is a group that is not a specific, concrete group such as $S(\square)$ or S_4.

For abstract groups, we use concise multiplicative notation where it will not cause confusion.

- We denote an abstract group simply by a single symbol such as G, without specifying a symbol for its binary operation.
- We denote a composite of two elements x and y of G simply by xy.

Warning: Unless the group is abelian, the composites xy and yx are not necessarily equal.

5. Theorems involving elements of order 2

Theorem B73

Let G be a group in which each element except the identity has order 2. Then G is abelian.

Theorem B74

Let G be a group of order greater than 2 in which each element except the identity has order 2. Then the order of G is a multiple of 4.

Theorem B75

Let G be a group of even order. Then G contains an element of order 2.

6. The **dihedral group** of order $2n$, for each integer $n \geq 3$, is the symmetry group of the regular polygon with n edges.

The **quaternion group** of order 8, denoted by Q_8, is a non-abelian group comprising the identity, six elements of order 4 and one element of order 2.

7. Isomorphism classes for groups of order up to 8

(See also the Quick reference section, page 137.)

Order	Standard group(s)	Properties
1	C_1	cyclic
2	C_2, $(\mathbb{Z}_2, +_2)$	cyclic
3	C_3, $(\mathbb{Z}_3, +_3)$	cyclic
4	C_4, $(\mathbb{Z}_4, +_4)$	cyclic
	V, $S(\square)$	abelian, non-cyclic
5	C_5, $(\mathbb{Z}_5, +_5)$	cyclic
6	C_6, $(\mathbb{Z}_6, +_6)$	cyclic
	$S(\triangle)$, D_3	non-abelian
7	C_7, $(\mathbb{Z}_7, +_7)$	cyclic
8	C_8, $(\mathbb{Z}_8, +_8)$	cyclic
	$S(\text{cuboid})$	abelian, non-cyclic
	$(U_{15}, \times_{15})$	abelian, non-cyclic
	$S(\square)$, D_4	non-abelian
	Q_8	non-abelian

Strategy B14 Groups of order 8

To determine the isomorphism class of a group of order 8, determine whether the group is abelian and find the number of elements of order 2. Then use the table in the Quick reference section on page 137.

3 Theorems and proofs in group theory

8. Intersection of subgroups

Theorem B81

Let H and K be subgroups of a group G. Then $H \cap K$ is also a subgroup of G.

Book C Linear algebra

Unit C1 Linear equations and matrices

1 Systems of linear equations

1. Linear equations

A linear equation has no terms that are products of unknowns or powers greater than 1.

A **linear equation** in the n **unknowns** $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

where $a_1, a_2, \ldots, a_n, b$ are real numbers, and $a_1, \ldots, a_n$ are not all zero. The numbers a_i are the **coefficients**, and b is the **constant term**.

A linear equation of the form

$$ax + by = c$$

represents a line in $\mathbb{R}^2$. There are infinitely many solutions to this equation – one corresponding to each point on the line.

A linear equation of the form

$$ax + by + cz = d$$

represents a plane in $\mathbb{R}^3$. There are infinitely many solutions to this equation – one corresponding to each point in the plane.

2. A **system of linear equations** is written

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + \cdots + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + \cdots + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + \cdots + & a_{mn}x_n & = & b_m. \end{array}$$

The **constant terms** are the numbers b_i, the **unknowns** are the variables x_j and the **coefficients** are the numbers a_{ij}. The coefficient of the jth unknown in the ith equation is a_{ij}.

The number of equations, m, need not be the same as the number of unknowns, n.

A **solution** of a system of m linear equations in n unknowns, $x_1, \ldots, x_n$, is a set of values $x_1 = c_1$, $x_2 = c_2, \ldots, x_n = c_n$ that simultaneously satisfy all m equations of the system.

The **solution set** of the system is the set of all the solutions.

3. A **consistent** system of linear equations has at least one solution, and an **inconsistent** system has no solutions.

4. A **homogeneous** system of linear equations is a system of linear equations in which each constant term is zero.

A **non-homogeneous** system contains at least one non-zero constant term.

The **trivial** solution to a homogeneous system of linear equations is the solution in which each unknown is equal to zero. Every homogeneous system has at least the trivial solution, so is consistent

A **non-trivial** solution has at least one unknown non-zero.

5. A system of m linear equations in n unknowns has a solution set which either

- contains exactly one solution
- is empty, or
- contains infinitely many solutions.

6. Elementary operations

The following operations do not change the solution set of a system of linear equations.

1. Interchange two equations.
2. Multiply an equation by a non-zero number.
3. Change one equation by adding to it a multiple of another.

7. Gauss–Jordan elimination is a systematic method for solving a system of linear equations by successively transforming the system into simpler systems without changing the solution set.

8. To solve a system of three linear equations in three unknowns, use elementary operations to try to reduce the system to the following form, where the asterisks denote numbers to be determined.

$$\begin{aligned} x &= * \\ y &= * \\ z &= * \end{aligned}$$

An equation of the form $0 = 0$ gives no constraints, whereas an equation of the form $0 = *$, where $* \neq 0$, implies the system is inconsistent.

2 Row-reduction

9. A **matrix** is a rectangular array of objects, called its **entries**, enclosed in brackets.

A **row** of a matrix comprises the entries along a horizontal line, and a **column** comprises those down a vertical line. A matrix of **size** $m \times n$ has m rows and n columns.

A **zero row** of a matrix is a row comprised entirely of zeros, and a **non-zero row** has at least one non-zero entry.

The **leading entry** of a row of a matrix is the first non-zero entry in the row (reading from left to right). A **leading 1** is a leading entry that is a 1.

10. The **augmented matrix** of a system of m linear equations in n unknowns $x_1, x_2, \ldots, x_n$ is

$$(\mathbf{A} \mid \mathbf{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{array}\right),$$

where the numbers b_i are the **constant terms** and a_{ij} are the **coefficients** of the system.

11. **Elementary row operations** are operations on the rows of the augmented matrix and correspond to elementary operations on the equations themselves, as follows.

1. Interchange two rows.
2. Multiply a row by a non-zero number.
3. Change one row by adding to it a multiple of another.

12. **Row-sum check**: To the right of each row write down a *check entry* equal to the sum of the entries in that row. Perform the row operations on these check entries also; they should remain the sum of the entries in the corresponding row. If this is not the case, then an error has been made.

13. A **row-reduced matrix** is a matrix satisfying the following four properties.

1. Any zero rows are at the bottom of the matrix.
2. Each non-zero row has a leading 1.
3. Each leading 1 is to the right of the leading 1 in the row above.
4. Each leading 1 is the only non-zero entry in its column.

$$\begin{pmatrix} 1 & * \cdots * & 0 & * \cdots * & 0 & * \cdots * & 0 & * \cdots * \\ & & 1 & * \cdots * & 0 & * \cdots * & 0 & * \cdots * \\ & & & & 1 & * \cdots * & 0 & * \cdots * \\ & 0 & & & & & 1 & * \cdots * \end{pmatrix}$$

14. **Unique solution** Whenever the original system of equations has a unique solution it can be written down directly from the row-reduced matrix.

No solution Whenever the original system of equations is inconsistent, row-reducing the augmented matrix yields a system that includes the equation $0 = 1$.

Infinitely many solutions Whenever the original system of equations has infinitely many solutions, the general solution can be determined by setting the non-leading unknowns equal to parameters and expressing all the unknowns in terms of these parameters.

15. **Row-reduction strategy**

Strategy C1 Row-reduction

To row-reduce a matrix using elementary row operations, carry out the following four steps, first with row 1 as the current row, then with row 2 as the current row, and so on, until

- **either** every row has been the current row
- **or** step 1 is not possible.

1. Select the first column from the left that has at least one non-zero entry in or below the current row.
2. If the current row has a 0 in the selected column, interchange it with a row below that has a non-zero entry in that column.
3. If the entry in the current row and the selected column is c, multiply the current row by $1/c$ to create a leading 1.
4. Add suitable multiples of the current row to the other rows to make each entry above and below the leading 1 into a 0.

You can use any elementary row operation when row-reducing an augmented matrix, but it is important not to use rows above the current row.

The different ways to row-reduce a matrix will always give the same answer.

Theorem C1

Every matrix has a unique row-reduced form.

16. Solving systems of linear equations using matrices

Strategy C2 Gauss–Jordan elimination

To use Gauss–Jordan elimination to solve a given system of linear equations:

1. form the augmented matrix
2. row-reduce the augmented matrix to obtain its row-reduced form (Strategy C1)
3. solve the simplified system of linear equations.

3 Matrix operations

17. A **square** matrix is an $n \times n$ matrix.

18. We write $\mathbf{A}$ or (a_{ij}) to denote a matrix:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = (a_{ij}).$$

The (i, j)**-entry** of a matrix $\mathbf{A}$, a_{ij}, is the entry in the ith row and jth column.

19. A **row vector** is a vector written with the components horizontally; such a vector can be regarded as a matrix with real entries that has just a single row.

A **column vector** is a vector with the components written vertically; such a vector can be regarded as a matrix with real entries that has just a single column.

It should be clear from the context whether an object is a vector in $\mathbb{R}^2$ or $\mathbb{R}^3$, with a geometrical interpretation, or a matrix with real entries.

20. $M_{m,n}$ denotes the set of all $m \times n$ matrices with real entries.

21. Matrix arithmetic

Equality Two $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ are **equal** if all their corresponding entries agree. We write $\mathbf{A} = \mathbf{B}$.

Zero matrix The $m \times n$ **zero matrix** $\mathbf{0}_{m,n}$ is the $m \times n$ matrix in which all entries are 0. It is denoted by $\mathbf{0}$ when it is clear from the context which size is intended.

Addition The **sum** of two $m \times n$ matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ is the $m \times n$ matrix $\mathbf{A} + \mathbf{B} = (a_{ij} + b_{ij})$ obtained by adding the corresponding entries:

$$\mathbf{A} + \mathbf{B} = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}.$$

Addition of matrices of different sizes is not defined.

Negatives The **negative** of an $m \times n$ matrix $\mathbf{A} = (a_{ij})$ is the $m \times n$ matrix $-\mathbf{A} = (-a_{ij})$ obtained by taking the negatives of its entries.

Subtraction The **difference** of two $m \times n$ matrices $\mathbf{A} = (a_{ij})$ and $\mathbf{B} = (b_{ij})$ is the $m \times n$ matrix $\mathbf{A} - \mathbf{B} = (a_{ij} - b_{ij})$ obtained by subtracting the corresponding entries:

$$\mathbf{A} - \mathbf{B} = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \cdots & a_{1n} - b_{1n} \\ a_{21} - b_{21} & a_{22} - b_{22} & \cdots & a_{2n} - b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} - b_{m1} & a_{m2} - b_{m2} & \cdots & a_{mn} - b_{mn} \end{pmatrix}.$$

Subtraction of matrices of different sizes is not defined.

Scalar multiplication The **scalar multiple** of an $m \times n$ matrix $\mathbf{A} = (a_{ij})$ by a real number k is the $m \times n$ matrix $k\mathbf{A}$ obtained by multiplying each entry in turn by k.

$$k\mathbf{A} = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix} = (ka_{ij}).$$

22. Addition in $M_{m,n}$

A1 Closure For all $\mathbf{A}, \mathbf{B} \in M_{m,n}$,

$$\mathbf{A} + \mathbf{B} \in M_{m,n}.$$

A2 Associativity For all $\mathbf{A}, \mathbf{B}, \mathbf{C} \in M_{m,n}$,

$$(\mathbf{A} + \mathbf{B}) + \mathbf{C} = \mathbf{A} + (\mathbf{B} + \mathbf{C}).$$

A3 Additive identity For all $\mathbf{A} \in M_{m,n}$ and $\mathbf{0} \in M_{m,n}$,

$$\mathbf{A} + \mathbf{0} = \mathbf{A} = \mathbf{0} + \mathbf{A}.$$

A4 Additive inverses For each $\mathbf{A} \in M_{m,n}$, there is a matrix $-\mathbf{A} \in M_{m,n}$ such that

$$\mathbf{A} + (-\mathbf{A}) = \mathbf{0} = -\mathbf{A} + \mathbf{A}.$$

A5 Commutativity For all $\mathbf{A}, \mathbf{B} \in M_{m,n}$,

$$\mathbf{A} + \mathbf{B} = \mathbf{B} + \mathbf{A}.$$

The **additive identity** in $M_{m,n}$ is $\mathbf{0}$ and the **additive inverse** of $\mathbf{A}$ in $M_{m,n}$ is $-\mathbf{A}$.

23. $(M_{m,n}, +)$ is an abelian group.

24. Combining addition and scalar multiplication in $M_{m,n}$

D1 Distributivity For all $\mathbf{A}, \mathbf{B} \in M_{m,n}$ and $k \in \mathbb{R}$,

$$k(\mathbf{A} + \mathbf{B}) = k\mathbf{A} + k\mathbf{B}.$$

25. Matrix multiplication

The **product** of an $m \times n$ matrix $\mathbf{A}$ with an $n \times p$ matrix $\mathbf{B}$ is the $m \times p$ matrix $\mathbf{AB}$ whose (i, j)-entry is obtained by multiplying each entry in the ith row of $\mathbf{A}$ by the corresponding entry in the jth column of $\mathbf{B}$ and adding the results.

In symbols, if $\mathbf{C} = \mathbf{AB}$, then

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}.$$

The product $\mathbf{AB}$ is not defined when the number of columns of the matrix $\mathbf{A}$ is not equal to the number of rows of the matrix $\mathbf{B}$.

26. Whenever these products can be formed, matrix multiplication is

- **associative**; that is, $(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC})$
- **distributive**; that is, $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC}$.

Matrix multiplication is **not commutative**: in general, $\mathbf{AB} \neq \mathbf{BA}$.

27. Diagonal matrices

The **diagonal** entries of a square matrix are the entries from the top left-hand corner to the bottom right-hand corner.

The **main diagonal** (or **leading diagonal**) of a square matrix comprises the diagonal entries of the matrix.

A **diagonal matrix** is a square matrix each of whose non-diagonal entries is zero.

The product of two diagonal matrices is another diagonal matrix, and its ith diagonal entry is the product of the ith diagonal entries of the matrices being multiplied.

Multiplication of diagonal matrices is **commutative**.

28. Positive **powers** of square matrices are defined as expected:

$$\mathbf{A}^2 = \mathbf{AA}, \quad \mathbf{A}^3 = \mathbf{AAA}, \quad \ldots.$$

29. An **upper triangular matrix** is a square matrix with each entry *below* the main diagonal equal to zero.

A **lower triangular matrix** is a square matrix with each entry *above* the main diagonal equal to zero.

A square row-reduced matrix is an upper triangular matrix.

A square matrix that is both upper triangular and lower triangular is a diagonal matrix.

30. The **identity matrix $\mathbf{I}_n$** is the $n \times n$ matrix

$$\begin{pmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

Each of the entries is 0 except those on the main diagonal, which are all 1.

It is denoted by $\mathbf{I}$ when the size is clear from the context.

Theorem C2

Let $\mathbf{A}$ be an $m \times n$ matrix. Then

$$\mathbf{I}_m\mathbf{A} = \mathbf{A} = \mathbf{AI}_n.$$

31. The **transpose** of an $m \times n$ matrix $\mathbf{A}$ is the $n \times m$ matrix $\mathbf{A}^T$ whose (i,j)-entry is the (j,i)-entry of $\mathbf{A}$. It is formed by interchanging the rows and columns of matrix $\mathbf{A}$.

Properties of matrix transposition

Let $\mathbf{A}$ and $\mathbf{B}$ be $m \times n$ matrices. Then:

1. $(\mathbf{A}^T)^T = \mathbf{A}$
2. $(\mathbf{A}+\mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T$.

Let $\mathbf{A}$ be an $m \times n$ matrix and $\mathbf{B}$ an $n \times p$ matrix. Then

3. $(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T$.

***Warning*:** Note the reversed order for the transpose of a product.

32. A square matrix $\mathbf{A}$ is **symmetric** if $\mathbf{A}^T = \mathbf{A}$.

33. The **matrix form** of a system of m linear equations in n unknowns $x_1, x_2, \dots, x_n$ is $\mathbf{Ax} = \mathbf{b}$:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

The **coefficient matrix** is the matrix $\mathbf{A}$ of coefficients of the system, the matrix of unknowns is $\mathbf{x}$, and the matrix of constant terms is $\mathbf{b}$.

4 Matrix inverses

34. Multiplication in $M_{n,n}$ (square matrices)

M1 Closure For all $\mathbf{A}, \mathbf{B} \in M_{n,n}$,

$$\mathbf{AB} \in M_{n,n}.$$

M2 Associativity For all $\mathbf{A}, \mathbf{B}, \mathbf{C} \in M_{n,n}$,

$$(\mathbf{AB})\mathbf{C} = \mathbf{A}(\mathbf{BC}).$$

M3 Multiplicative identity For all $\mathbf{A} \in M_{n,n}$,

$$\mathbf{AI}_n = \mathbf{A} = \mathbf{I}_n\mathbf{A}.$$

The **multiplicative identity** in $M_{n,n}$ is $\mathbf{I}_n$.

35. Combining addition and multiplication in $M_{n,n}$

D1 Distributivity For all $\mathbf{A}, \mathbf{B}, \mathbf{C} \in M_{n,n}$,

$$\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC},$$

and

$$(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}.$$

36. Multiplicative inverse

The matrix $\mathbf{B}$ is a **multiplicative inverse** (or **inverse**) of $\mathbf{A}$ in $M_{n,n}$ if $\mathbf{A}, \mathbf{B} \in M_{n,n}$ and

$$\mathbf{AB} = \mathbf{I}_n = \mathbf{BA}.$$

Theorem C3

If a square matrix has an inverse, then this inverse is unique.

Theorem C4

A square matrix with a zero row has no inverse.

37. An **invertible** matrix is a square matrix that has an inverse. The unique inverse of an invertible matrix $\mathbf{A}$ is denoted by $\mathbf{A}^{-1}$. For any invertible matrix $\mathbf{A}$,

$$\mathbf{AA}^{-1} = \mathbf{I} \quad \text{and} \quad \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}.$$

If $\mathbf{A}$ is an invertible matrix, then $\mathbf{A}^{-1}$ is also invertible, with inverse $\mathbf{A}$; that is,

$$(\mathbf{A}^{-1})^{-1} = \mathbf{A}.$$

Let $\mathbf{A}$ and $\mathbf{B}$ be invertible matrices of the same size. Then $\mathbf{AB}$ is invertible, and

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}.$$

***Warning*:** Note the reversed order here.

Theorem C5

Let $\mathbf{A}_1, \mathbf{A}_2, \dots, \mathbf{A}_k$ be invertible matrices of the same size. Then the product $\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k$ is invertible, with

$$(\mathbf{A}_1\mathbf{A}_2\cdots\mathbf{A}_k)^{-1} = \mathbf{A}_k^{-1}\mathbf{A}_{k-1}^{-1}\cdots\mathbf{A}_1^{-1}.$$

Theorem C6

The set of all invertible $n \times n$ matrices with real entries forms a group under matrix multiplication.

38. Inverting matrices

Theorem C7 Invertibility Theorem

(a) A square matrix is invertible if and only if its row-reduced form is $\mathbf{I}$.

(b) Any sequence of elementary row operations that transforms a matrix $\mathbf{A}$ to $\mathbf{I}$ also transforms $\mathbf{I}$ to $\mathbf{A}^{-1}$.

Strategy C3 Matrix inversion

To determine whether or not a given square matrix $\mathbf{A}$ is invertible, and to find its inverse if it is, do the following.

Write down $(\mathbf{A} \mid \mathbf{I})$, and row-reduce it (Strategy C1) until the left half is in row-reduced form.

- If the left half is the identity matrix, then the right half is $\mathbf{A}^{-1}$.
- Otherwise, $\mathbf{A}$ is not invertible.

You may find it helpful to remember the following scheme for this strategy:

$$\begin{array}{c} (\mathbf{A} \mid \mathbf{I}) \\ \downarrow \\ (\mathbf{I} \mid \mathbf{A}^{-1}). \end{array}$$

If it becomes clear while you are row-reducing $(\mathbf{A} \mid \mathbf{I})$ that the left half will not reduce to the identity matrix (for example, if a zero row appears in the left half), then you can immediately conclude that $\mathbf{A}$ is not invertible.

39. Invertibility and systems of linear equations

Theorem C8

Let $\mathbf{A}$ be an invertible matrix. Then the system of linear equations $\mathbf{A}\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$.

Theorem C9

Let $\mathbf{A}$ be an $n \times n$ matrix. Then the following statements are equivalent.

(a) $\mathbf{A}$ is invertible.

(b) The system $\mathbf{A}\mathbf{x} = \mathbf{b}$ has a unique solution for each $n \times 1$ matrix $\mathbf{b}$.

(c) The system $\mathbf{A}\mathbf{x} = \mathbf{0}$ has only the trivial solution.

40. An **elementary matrix** is a matrix obtained by performing an elementary row operation on an identity matrix.

Theorem C10

Let $\mathbf{E}$ be an elementary matrix, and let $\mathbf{A}$ be any matrix with the same number of rows as $\mathbf{E}$. Then the product $\mathbf{E}\mathbf{A}$ is the same as the matrix obtained when the elementary row operation associated with $\mathbf{E}$ is performed on $\mathbf{A}$.

Corollary C11

Let $\mathbf{E}_1, \mathbf{E}_2, \dots, \mathbf{E}_k$ be the $m \times m$ elementary matrices associated with a sequence of k elementary row operations carried out on a matrix $\mathbf{A}$ with m rows, in the same order. Then, after the sequence of row operations has been performed, the resulting matrix is

$$\mathbf{E}_k\mathbf{E}_{k-1} \cdots \mathbf{E}_2\mathbf{E}_1\mathbf{A}.$$

41. Inverse elementary row operations

Given any elementary row operation, the *inverse* elementary row operation that undoes the effect of the first, can be found, as summarised in the following table.

Elementary row operation	Inverse elementary row operation
$\mathbf{r}_i \leftrightarrow \mathbf{r}_j$	$\mathbf{r}_i \leftrightarrow \mathbf{r}_j$
$\mathbf{r}_i \to c\,\mathbf{r}_i \quad (c \neq 0)$	$\mathbf{r}_i \to (1/c)\,\mathbf{r}_i$
$\mathbf{r}_i \to \mathbf{r}_i + c\,\mathbf{r}_j$	$\mathbf{r}_i \to \mathbf{r}_i - c\,\mathbf{r}_j$

> **Theorem C12**
>
> Let $\mathbf{E}_1$ and $\mathbf{E}_2$ be elementary matrices of the same size whose associated elementary row operations are inverses of each other. Then $\mathbf{E}_1$ and $\mathbf{E}_2$ are inverses of each other.
>
> **Corollary C13**
>
> Every elementary matrix is invertible, and its inverse is also an elementary matrix.

5 Determinants

42. The **determinant** of a 2×2 matrix

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

is

$$\det \mathbf{A} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

> **Strategy C4 Inverse of a 2×2 matrix**
>
> To find the inverse of a 2×2 matrix
>
> $$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$
>
> with $\det \mathbf{A} = ad - bc \neq 0$:
>
> - interchange the diagonal entries
> - multiply the non-diagonal entries by -1
> - divide by the determinant of $\mathbf{A}$, giving
>
> $$\mathbf{A}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

43. The **determinant** of a 3×3 matrix

$$\mathbf{A} = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$$

is

$$\det \mathbf{A} = a_1 \begin{vmatrix} b_2 & c_2 \\ b_3 & c_3 \end{vmatrix} - b_1 \begin{vmatrix} a_2 & c_2 \\ a_3 & c_3 \end{vmatrix} + c_1 \begin{vmatrix} a_2 & b_2 \\ a_3 & b_3 \end{vmatrix}.$$

Notice the minus sign before the second term.

44. A **submatrix** is a matrix formed from another matrix with some of the rows and/or columns removed.

45. The **cofactor** A_{ij} associated with the entry a_{ij} of an $n \times n$ matrix $\mathbf{A} = (a_{ij})$ is

$$A_{ij} = (-1)^{i+j} \det \mathbf{A}_{ij},$$

where $\mathbf{A}_{ij}$ is the $(n-1) \times (n-1)$ submatrix of $\mathbf{A}$ resulting when the ith row and jth column (the row and column containing the entry a_{ij}) are removed.

For example, for a 3×3 matrix $\mathbf{A} = (a_{ij})$, to find the cofactor A_{12} we have

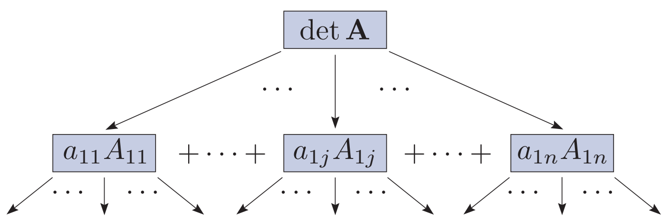
$$(-1)^{1+2} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \quad \text{so } A_{12} = -\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix}.$$

46. The **determinant** of an $n \times n$ matrix $\mathbf{A} = (a_{ij})$ is

$$\det \mathbf{A} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$
$$= a_{11}A_{11} + a_{12}A_{12} + \cdots + a_{1n}A_{1n}.$$

Warning: Do not forget the minus sign that is a part of alternate cofactors!

> **Strategy C5 Finding the determinant**
>
> To evaluate the determinant of an $n \times n$ matrix:
>
> 1. expand along the top row to express the $n \times n$ determinant in terms of n determinants of size $(n-1) \times (n-1)$
> 2. expand along the top row of each of the resulting determinants
> 3. repeatedly apply step 2 until the only determinants in the expression are of size 2×2
> 4. evaluate the final expression.



47. Properties of determinants

Theorem C14

Let $\mathbf{A}$ and $\mathbf{B}$ be two square matrices of the same size. Then the following hold:

(a) $\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B})$

(b) $\det \mathbf{I} = 1$

(c) $\det \mathbf{A}^T = \det \mathbf{A}$.

48. Determinants and elementary matrices

Theorem C15

Let $\mathbf{E}$ be an elementary matrix, and let k be a non-zero real number.

(a) If $\mathbf{E}$ results from interchanging two rows of $\mathbf{I}$, then $\det \mathbf{E} = -1$.

(b) If $\mathbf{E}$ results from multiplying a row of $\mathbf{I}$ by k, then $\det \mathbf{E} = k$.

(c) If $\mathbf{E}$ results from adding k times one row of $\mathbf{I}$ to another row, then $\det \mathbf{E} = 1$.

49. Matrices with zero determinant

Theorem C16

Let $\mathbf{A}$ be a square matrix. Then $\det \mathbf{A} = 0$ if any of the following hold:

(a) $\mathbf{A}$ has an entire row (or column) of zeros

(b) $\mathbf{A}$ has two equal rows (or columns)

(c) $\mathbf{A}$ has two proportional rows (or columns).

50. Determinants and inverses of matrices

Theorem C17

A square matrix $\mathbf{A}$ is invertible if and only if $\det \mathbf{A} \neq 0$.

Theorem C18

Let $\mathbf{A}$ and $\mathbf{B}$ be square matrices of the same size. Then $\mathbf{AB} = \mathbf{I}$ if and only if $\mathbf{BA} = \mathbf{I}$.

51. The Summary Theorem collects together Theorems C7, C9 and C17.

Theorem C19 Summary Theorem

Let $\mathbf{A}$ be an $n \times n$ matrix. Then the following statements are equivalent.

(a) $\mathbf{A}$ is invertible.

(b) $\det \mathbf{A} \neq 0$.

(c) The row-reduced form of $\mathbf{A}$ is $\mathbf{I}_n$.

(d) The system $\mathbf{Ax} = \mathbf{b}$ has precisely one solution for each $n \times 1$ matrix $\mathbf{b}$.

(e) The system $\mathbf{Ax} = \mathbf{0}$ has only the trivial solution.

52. Summary of properties of matrices

Let $\mathbf{A}$ and $\mathbf{B}$ be two square matrices of the same size. Then

$$\det(\mathbf{AB}) = (\det \mathbf{A})(\det \mathbf{B}),$$

$$(\mathbf{AB})^T = \mathbf{B}^T\mathbf{A}^T,$$

$$\det \mathbf{A}^T = \det \mathbf{A}.$$

If $\mathbf{A}$ and $\mathbf{B}$ are invertible, then

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1},$$

$$\det \mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}}.$$

Unit C2 Vector spaces

1 Vector spaces

1. In $\mathbb{R}^2$, the set of ordered pairs of real numbers, the operations of **addition** and of **multiplication by a scalar** are defined as:

$$(u_1, u_2) + (v_1, v_2) = (u_1 + v_1, u_2 + v_2),$$
$$\alpha(u_1, u_2) = (\alpha u_1, \alpha u_2), \quad \text{where } \alpha \in \mathbb{R}.$$

In $\mathbb{R}^3$, the set of ordered triples of real numbers, the operations of **addition** and of **multiplication by a scalar** are defined as:

$$\begin{aligned}(u_1, u_2, u_3) &+ (v_1, v_2, v_3)\\ &= (u_1 + v_1, u_2 + v_2, u_3 + v_3),\end{aligned}$$
$$\alpha(u_1, u_2, u_3) = (\alpha u_1, \alpha u_2, \alpha u_3), \quad \text{where } \alpha \in \mathbb{R}.$$

2. An **ordered n-tuple**, for a positive integer n, is a sequence of real numbers $(u_1, u_2, \dots, u_n)$.

n-dimensional space, $\mathbb{R}^n$, is the set of all ordered n-tuples.

The space $\mathbb{R}^n$ is often called a **Euclidean space** and its elements $(u_1, u_2, \dots, u_n)$ are called vectors.

3. Let

$$\mathbf{u} = (u_1, u_2, \dots, u_n) \quad \text{and} \quad \mathbf{v} = (v_1, v_2, \dots, v_n)$$

be two vectors in $\mathbb{R}^n$. The operations of **addition** and of **multiplication by a scalar** are defined as:

$$\begin{aligned}\mathbf{u} + \mathbf{v} &= (u_1, u_2, \dots, u_n) + (v_1, v_2, \dots, v_n)\\ &= (u_1 + v_1, u_2 + v_2, \dots, u_n + v_n),\\ \alpha\mathbf{u} &= (\alpha u_1, \alpha u_2, \dots, \alpha u_n), \quad \text{where } \alpha \in \mathbb{R}.\end{aligned}$$

4. Vector space definition

A (**real**) **vector space** consists of a set V of elements called **vectors** and two operations, vector addition and scalar multiplication (by a real number), such that the following axioms hold.

Axioms for addition

A1 Closure For all $\mathbf{v}_1, \mathbf{v}_2 \in V$,

$$\mathbf{v}_1 + \mathbf{v}_2 \in V.$$

A2 Associativity For all $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3 \in V$,

$$(\mathbf{v}_1 + \mathbf{v}_2) + \mathbf{v}_3 = \mathbf{v}_1 + (\mathbf{v}_2 + \mathbf{v}_3).$$

A3 Additive identity For all $\mathbf{v} \in V$, there is a zero element $\mathbf{0} \in V$ satisfying

$$\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}.$$

A4 Additive inverses For each $\mathbf{v} \in V$, there is an element $-\mathbf{v}$ (its additive inverse) such that

$$\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = -\mathbf{v} + \mathbf{v}.$$

A5 Commutativity For all $\mathbf{v}_1, \mathbf{v}_2 \in V$,

$$\mathbf{v}_1 + \mathbf{v}_2 = \mathbf{v}_2 + \mathbf{v}_1.$$

Axioms for scalar multiplication

S1 Closure For all $\mathbf{v} \in V$, and $\alpha \in \mathbb{R}$,

$$\alpha\mathbf{v} \in V.$$

S2 Associativity For all $\mathbf{v} \in V$, and $\alpha, \beta \in \mathbb{R}$,

$$\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}.$$

S3 Scalar multiplicative identity For all $\mathbf{v} \in V$,

$$1\mathbf{v} = \mathbf{v}.$$

Axioms combining addition and scalar multiplication

D1 Distributivity For all $\mathbf{v}_1, \mathbf{v}_2 \in V$, and $\alpha \in \mathbb{R}$,

$$\alpha(\mathbf{v}_1 + \mathbf{v}_2) = \alpha\mathbf{v}_1 + \alpha\mathbf{v}_2.$$

D2 Distributivity For all $\mathbf{v} \in V$, and $\alpha, \beta \in \mathbb{R}$,

$$(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}.$$

5. Axioms A1–A5 imply that $(V, +)$ is an **abelian group**.

6. Examples of vector spaces

- $\mathbb{R}$, $\mathbb{R}^2$, $\mathbb{R}^3$, and (more generally) $\mathbb{R}^n$ ($n \geq 1$).
- $\mathbb{R}^\infty$, the set of all infinite sequences of real numbers.
- $\mathbb{C}$, the set of complex numbers.
- $M_{m,n}$, the set of all $m \times n$ matrices with real entries.
- P_n, the set of all real polynomials of degree less than n.
- $V = \{a\cos x + b\sin x : a, b \in \mathbb{R}\}$.

2 Linear combinations and spanning sets

7. A **linear combination** of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ in a vector space V, is a vector of the form

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are real numbers. This vector also belongs to V.

8. Spanning sets

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be a finite set of vectors in a vector space V.

The **span** $\langle S \rangle$ of S is the set of all possible linear combinations

$$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k,$$

where $\alpha_1, \alpha_2, \ldots, \alpha_k$ are real numbers; that is,

$$\langle S \rangle = \{\alpha_1\mathbf{v}_1 + \cdots + \alpha_k\mathbf{v}_k : \alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}\}.$$

The set of vectors $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ **spans** $\langle S \rangle$; it is a **spanning set** for $\langle S \rangle$, and $\langle S \rangle$ is the set **spanned** by S.

9. Linear combinations and spanning sets

Strategy C6 Linear combinations

To determine whether a given vector $\mathbf{v}$ can be written as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$:

1. write $\mathbf{v} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k$
2. use this expression to write down a system of linear equations in the unknowns $\alpha_1, \alpha_2, \ldots, \alpha_k$
3. solve the resulting system of equations, if possible.

Then $\mathbf{v}$ can be written as a linear combination of $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k$ if and only if the system has a solution.

Strategy C6 can be used to show that a given set of vectors is a spanning set for the whole of a vector space.

***Warning*:** When the strategy results in more equations than unknowns it is important to confirm the consistency of the system by checking *all* the equations.

3 Bases and dimension

10. A **minimal spanning set** of a vector space V is a set containing the smallest number of vectors that span V.

Theorem C20

Suppose that the vector $\mathbf{v}_k$ can be written as a linear combination of the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}$. Then the span of the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is the same as the span of the set $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{k-1}\}$.

11. Linear dependence and independence

A finite set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ in a vector space V is

- **linearly dependent** if there exist real numbers $\alpha_1, \alpha_2, \ldots, \alpha_k$, *not all zero*, such that
  $$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0}$$
- **linearly independent** if it is not linearly dependent; that is, if
  $$\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0}$$
  implies $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$.

Strategy C7 Linear independence

To test whether a given set of vectors $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ is linearly independent:

1. write down the equation $\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_k\mathbf{v}_k = \mathbf{0}$
2. express this equation as a system of linear equations in the unknowns $\alpha_1, \alpha_2, \ldots, \alpha_k$
3. solve the resulting system of equations.

If the only solution is $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 0$, then the set of vectors is linearly independent.

If there is a solution with at least one of $\alpha_1, \alpha_2, \ldots, \alpha_k$ not equal to zero, then the set of vectors is linearly dependent.

- A linearly independent set cannot contain $\mathbf{0}$.

The following are linearly independent sets:

- one non-zero vector
- two non-zero vectors if and only if the vectors are not multiples of one another.

12. A **basis** for a vector space V is a minimal spanning set. A basis is not unique, so V can have many different bases.

A set is a minimum spanning set if and only if it is linearly independent.

Theorem C21

Let S be a basis for a vector space V. Then each vector in V can be expressed as a linear combination of the vectors in S in only one way.

Strategy C8 Checking a basis – span

To determine whether a set of vectors S in a vector space V is a basis for V, check the following conditions.

(1) S is linearly independent (Strategy C7).

(2) S spans V (Strategy C6).

If both (1) and (2) hold, then S is a basis for V.

If either (1) or (2) does not hold, then S is not a basis for V.

13. The $\boldsymbol{E}$-coordinate representation of v

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a basis for a vector space V, and suppose that

$$\mathbf{v} = v_1\mathbf{e}_1 + v_2\mathbf{e}_2 + \cdots + v_n\mathbf{e}_n,$$

where $v_1, v_2, \ldots, v_n \in \mathbb{R}$.

The **$\boldsymbol{E}$-coordinate representation** of $\mathbf{v}$ is

$$\mathbf{v}_E = (v_1, v_2, \ldots, v_n)_E.$$

The **$\boldsymbol{E}$-coordinates** of $\mathbf{v}$ (the **coordinates of v with respect to the basis $\boldsymbol{E}$**) are $v_1, v_2, \ldots, v_n$.

If E is a standard basis, then we refer to the *standard coordinate representation*, *standard coordinates*, and so on; in this case, the subscript E is usually omitted.

14. Let V be a vector space:

- V is **finite-dimensional** if it contains a finite set of vectors S that forms a basis for V
- V is **infinite-dimensional** if no such set exists.

15. The Basis Theorem

Theorem C22

Let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be a basis for a vector space V, and let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be a set of m vectors in V, where $m > n$. Then S is a linearly dependent set.

Corollary C23

Let V be a vector space with a basis containing n vectors. If a linearly independent subset of V contains m vectors, then $m \leq n$.

Theorem C24 Basis Theorem

Let V be a finite-dimensional vector space. Then every basis for V contains the same number of vectors.

16. Dimension of a vector space

The **dimension** $\dim V$ of a finite-dimensional vector space V is the number of vectors in any basis for the space.

Theorem C25

Let V be an n-dimensional vector space. Then any set of n linearly independent vectors in V is a basis for V.

Strategy C9 Checking a basis – dim

To determine whether a set of vectors S in a vector space V of dimension n is a basis for V, check the following conditions.

(1) S contains n vectors.

(2) S is linearly independent (Strategy C7).

If both (1) and (2) hold, then S is a basis for V.

If either (1) or (2) does not hold, then S is not a basis for V.

17. Any linearly independent set can be extended to form a basis.

> **Theorem C26**
>
> Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be a linearly independent subset of an n-dimensional vector space V, where $m < n$. Then there exist vectors $\mathbf{v}_{m+1}, \ldots, \mathbf{v}_n$ in V such that $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is a basis for V.

18. Standard bases

$\mathbb{R}^2 :\ \{(1,0),(0,1)\}$

$\mathbb{R}^3 :\ \{(1,0,0),(0,1,0),(0,0,1)\}$

$\mathbb{R}^n :\ \{(1,0,\ldots,0),(0,1,0,\ldots,0),\ldots,(0,\ldots,0,1)\}$

$\mathbb{C} :\ \{1,i\}$

$M_{2,2} :\ \left\{\begin{pmatrix}1&0\\0&0\end{pmatrix},\begin{pmatrix}0&1\\0&0\end{pmatrix},\begin{pmatrix}0&0\\1&0\end{pmatrix},\begin{pmatrix}0&0\\0&1\end{pmatrix}\right\}$

$P_n :\ \{1,x,x^2,\ldots,x^{n-1}\}$

19. Dimensions of familiar vector spaces

The dimension of

- $\mathbb{R}^n$ is n, so $\dim R^2 = 2$ and $\dim \mathbb{R}^3 = 3$
- $\mathbb{C}$ is 2
- P_n is n, so $\dim P_2 = 2$ and $\dim P_3 = 3$
- $M_{m,n}$ is $m \times n$, so $\dim M_{2,2} = 4$.

4 Subspaces

20. A **subspace** of a vector space V is a subset S of V that is itself a vector space under vector addition and scalar multiplication as defined in V.

> **Theorem C27**
>
> A subset S of a vector space V is a subspace of V if it satisfies the following conditions.
>
> (a) $\mathbf{0} \in S$.
>
> (b) S is closed under vector addition.
>
> (c) S is closed under scalar multiplication.

> **Theorem C28**
>
> Let S be a non-empty finite subset of a vector space V. Then $\langle S\rangle$ is a subspace of V.

> **Theorem C29**
>
> The dimension of a subspace of a vector space V is less than or equal to the dimension of V.

> **Strategy C10 Subspace**
>
> To test whether a given subset S of a vector space V is a subspace of V, check the following conditions.
>
> (1) $\mathbf{0} \in S$ (zero vector).
>
> (2) If $\mathbf{v}_1, \mathbf{v}_2 \in S$, then $\mathbf{v}_1 + \mathbf{v}_2 \in S$ (vector addition).
>
> (3) If $\mathbf{v} \in S$ and $\alpha \in \mathbb{R}$, then $\alpha\mathbf{v} \in S$ (scalar multiplication).
>
> If (1), (2) and (3) hold, then S is a subspace of V.
>
> If any of (1), (2) or (3) does not hold, then S is not a subspace of V.

21. Basis of a subspace

When using Strategy C8 or Strategy C9 to check whether a set of vectors S is a basis of a *proper* subspace T of V, you must also check that *all* the vectors in S are contained in the subspace T.

5 Orthogonal bases

22. Orthogonality in $\mathbb{R}^3$

The vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in $\mathbb{R}^3$ are **orthogonal** if $\mathbf{v}_1 \cdot \mathbf{v}_2 = 0$.

An **orthogonal set** of vectors in $\mathbb{R}^3$ is a set of vectors in $\mathbb{R}^3$ where every pair of distinct vectors in the set is orthogonal.

An **orthogonal basis** for $\mathbb{R}^3$ is an orthogonal set that is a basis for $\mathbb{R}^3$.

> **Theorem C30**
>
> Let $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ be an orthogonal set of non-zero vectors in $\mathbb{R}^3$. Then $\mathbf{v}_1$, $\mathbf{v}_2$ and $\mathbf{v}_3$ are linearly independent.

Theorem C31

Any orthogonal set of three non-zero vectors in $\mathbb{R}^3$ is an orthogonal basis for $\mathbb{R}^3$.

Strategy C11 Orthogonal basis in $\mathbb{R}^3$

To express a vector $\mathbf{u}$ in $\mathbb{R}^3$ in terms of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$:

1. calculate $\alpha_1 = \dfrac{\mathbf{v}_1 \cdot \mathbf{u}}{\mathbf{v}_1 \cdot \mathbf{v}_1}$, $\alpha_2 = \dfrac{\mathbf{v}_2 \cdot \mathbf{u}}{\mathbf{v}_2 \cdot \mathbf{v}_2}$ and $\alpha_3 = \dfrac{\mathbf{v}_3 \cdot \mathbf{u}}{\mathbf{v}_3 \cdot \mathbf{v}_3}$
2. write $\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \alpha_3\mathbf{v}_3$.

23. Orthogonality in $\mathbb{R}^n$

Let $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ and $\mathbf{w} = (w_1, w_2, \ldots, w_n)$ be vectors in $\mathbb{R}^n$.

The **scalar product** of $\mathbf{v}$ and $\mathbf{w}$ is the real number

$$\mathbf{v} \cdot \mathbf{w} = v_1w_1 + v_2w_2 + \cdots + v_nw_n.$$

The vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^n$ are **orthogonal** if $\mathbf{v} \cdot \mathbf{w} = 0$.

An **orthogonal set** of vectors in $\mathbb{R}^n$ is a set of vectors in $\mathbb{R}^n$ where every pair of distinct vectors in the set is orthogonal.

An **orthogonal basis** for $\mathbb{R}^n$ is an orthogonal set that is a basis for $\mathbb{R}^n$.

Theorem C32

Let $S = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_k\}$ be an orthogonal set of non-zero vectors in $\mathbb{R}^n$. Then S is a linearly independent set.

Corollary C33

Any orthogonal set of n non-zero vectors in $\mathbb{R}^n$ is an orthogonal basis for $\mathbb{R}^n$.

In $\mathbb{R}^2$ and $\mathbb{R}^3$ the term **perpendicular** is also used to describe orthogonal vectors.

24. Expressing vectors in terms of orthogonal bases

Theorem C34

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an orthogonal basis for $\mathbb{R}^n$ and let $\mathbf{u}$ be any vector in $\mathbb{R}^n$. Then

$$\mathbf{u} = \left(\frac{\mathbf{v}_1 \cdot \mathbf{u}}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 + \left(\frac{\mathbf{v}_2 \cdot \mathbf{u}}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2 + \cdots + \left(\frac{\mathbf{v}_n \cdot \mathbf{u}}{\mathbf{v}_n \cdot \mathbf{v}_n}\right)\mathbf{v}_n.$$

Strategy C12 Orthogonal basis $\mathbb{R}^n$

To express a vector $\mathbf{u}$ in $\mathbb{R}^n$ in terms of an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$:

1. calculate
$$\alpha_1 = \frac{\mathbf{v}_1 \cdot \mathbf{u}}{\mathbf{v}_1 \cdot \mathbf{v}_1},\ \alpha_2 = \frac{\mathbf{v}_2 \cdot \mathbf{u}}{\mathbf{v}_2 \cdot \mathbf{v}_2},\ \ldots,\ \alpha_n = \frac{\mathbf{v}_n \cdot \mathbf{u}}{\mathbf{v}_n \cdot \mathbf{v}_n}$$
2. write $\mathbf{u} = \alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n$.

25. Obtaining an orthogonal basis in $\mathbb{R}^n$ from an arbitrary basis

Theorem C35 Gram–Schmidt orthogonalisation

Let $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ be a basis for $\mathbb{R}^n$, and let

$$\begin{aligned}
\mathbf{v}_1 &= \mathbf{w}_1,\\
\mathbf{v}_2 &= \mathbf{w}_2 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{w}_2}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1,\\
\mathbf{v}_3 &= \mathbf{w}_3 - \left(\frac{\mathbf{v}_1 \cdot \mathbf{w}_3}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{w}_3}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2,\\
&\ \vdots\\
\mathbf{v}_n &= \mathbf{w}_n - \left(\frac{\mathbf{v}_1 \cdot \mathbf{w}_n}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right)\mathbf{v}_1 - \left(\frac{\mathbf{v}_2 \cdot \mathbf{w}_n}{\mathbf{v}_2 \cdot \mathbf{v}_2}\right)\mathbf{v}_2\\
&\quad - \cdots - \left(\frac{\mathbf{v}_{n-1} \cdot \mathbf{w}_n}{\mathbf{v}_{n-1} \cdot \mathbf{v}_{n-1}}\right)\mathbf{v}_{n-1}.
\end{aligned}$$

Then $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ is an orthogonal basis for $\mathbb{R}^n$.

26. To find an orthogonal basis of a *proper* subspace T of $\mathbb{R}^n$ of dimension m given a basis $S = \{\mathbf{w}_1, \mathbf{w}_2, \dots \mathbf{w}_m\}$, use Gram–Schmidt orthogonalisation on S to obtain an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots \mathbf{v}_m\}$.

27. To find an orthogonal basis for $\mathbb{R}^3$ containing a given vector $\mathbf{v}$, find an orthogonal basis for the plane through the origin that has normal vector $\mathbf{v}$ by:

1. using the vector equation of a plane $\mathbf{x} \cdot \mathbf{v} = 0$ to find any pair of linearly independent vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ in the plane
2. using the Gram–Schmidt orthogonalisation process on $\mathbf{w}_1$ and $\mathbf{w}_2$ to find vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ forming an orthogonal basis for this plane.

The set $\{\mathbf{v}, \mathbf{v}_1, \mathbf{v}_2\}$ is an orthogonal basis for $\mathbb{R}^3$.

28. The **magnitude** of a vector $\mathbf{v} = (v_1, v_2, \dots, v_n)$ in $\mathbb{R}^n$ is

$$|\mathbf{v}| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = \sqrt{v_1^2 + v_2^2 + \cdots + v_n^2}.$$

29. An **orthonormal basis** for $\mathbb{R}^n$ is an orthogonal basis in which each basis vector has magnitude 1.

Strategy C13 Orthonormal basis $\mathbb{R}^n$

To construct an orthonormal basis for $\mathbb{R}^n$ from an orthogonal basis $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ for $\mathbb{R}^n$:

1. calculate the magnitude of each basis vector
2. scalar multiply each basis vector by the reciprocal of its magnitude.

The required orthonormal basis is

$$\left\{ \frac{\mathbf{v}_1}{|\mathbf{v}_1|}, \frac{\mathbf{v}_2}{|\mathbf{v}_2|}, \dots, \frac{\mathbf{v}_n}{|\mathbf{v}_n|} \right\}.$$

Theorem C36

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be an orthonormal basis for $\mathbb{R}^n$, and let $\mathbf{u}$ be any vector in $\mathbb{R}^n$. Then

$$\mathbf{u} = (\mathbf{v}_1 \cdot \mathbf{u})\mathbf{v}_1 + (\mathbf{v}_2 \cdot \mathbf{u})\mathbf{v}_2 + \cdots + (\mathbf{v}_n \cdot \mathbf{u})\mathbf{v}_n.$$

Unit C3 Linear transformations

Throughout, V and W denote vector spaces.

1 Introducing linear transformations

1. A **linear transformation of the plane** is a function from $\mathbb{R}^2$ to $\mathbb{R}^2$ that:

- maps parallel lines to parallel lines
- preserves scalar multiples
- maps the zero vector to itself.

2. The following are linear transformations of the plane, for any real numbers k and l.

A k-**dilation** of $\mathbb{R}^2$ scales (or stretches) vectors by a factor k with respect to the origin.

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ky \end{pmatrix}$$

A (k,l)-**scaling** of $\mathbb{R}^2$ scales vectors by a factor k in the x-direction and by a factor l in the y-direction.

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} k & 0 \\ 0 & l \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} kx \\ ly \end{pmatrix}$$

A **rotation** r_θ of $\mathbb{R}^2$ rotates vectors anticlockwise through an angle θ about the origin.

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\theta - y\sin\theta \\ x\sin\theta + y\cos\theta \end{pmatrix}$$

A **reflection** q_ϕ of $\mathbb{R}^2$ reflects vectors in the straight line through the origin that makes an angle ϕ with the positive x-axis (measured anticlockwise).

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} \cos 2\phi & \sin 2\phi \\ \sin 2\phi & -\cos 2\phi \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos 2\phi + y\sin 2\phi \\ x\sin 2\phi - y\cos 2\phi \end{pmatrix}.$$

A **shear** of $\mathbb{R}^2$ in the x-direction by a factor k scales (or stretches) vectors in the x-direction by a factor of k with respect to the origin.

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + ky \\ y \end{pmatrix}$$

3. A **translation** of $\mathbb{R}^2$ by (a, b) is the function

$$\begin{aligned} t : \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ (x, y) &\longmapsto (x + a, y + b). \end{aligned}$$

A translation is not a linear transformation unless $a = b = 0$.

4. A **linear transformation** from a vector space V to a vector space W is a function $t : V \longrightarrow W$ that satisfies properties LT1 and LT2.

LT1 $t(\mathbf{v}_1 + \mathbf{v}_2) = t(\mathbf{v}_1) + t(\mathbf{v}_2)$, for all $\mathbf{v}_1, \mathbf{v}_2 \in V$.

LT2 $t(\alpha\mathbf{v}) = \alpha\, t(\mathbf{v})$, for all $\mathbf{v} \in V$, $\alpha \in \mathbb{R}$.

Theorem C37

Let $t : V \longrightarrow W$ be a linear transformation. Then $t(\mathbf{0}) = \mathbf{0}$.

Strategy C14 Linear transformation

To determine whether or not a given function $t : V \longrightarrow W$ is a linear transformation, do the following.

1. Check whether $t(\mathbf{0}) = \mathbf{0}$; if not, then t is not a linear transformation.
2. Check whether t satisfies the two properties LT1 and LT2.

The function t is a linear transformation if and only if both these properties are satisfied.

Only step 2 of Strategy C14 is required to *show* that a function t is a linear transformation.

5. Definition combining LT1 and LT2

A linear transformation preserves linear combinations of vectors.

Theorem C38

A function $t : V \longrightarrow W$ is a linear transformation if and only if it satisfies

LT3 $t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2) = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2)$, for all $\mathbf{v}_1, \mathbf{v}_2 \in V$ and all $\alpha_1, \alpha_2 \in \mathbb{R}$.

6. Linear combinations of vectors

Given a linear transformation $t : V \longrightarrow W$ and the images of the basis vectors, we can determine the image, in W, of any vector in V.

Theorem C39

Let $t : V \longrightarrow W$ be a linear transformation. Then

$$\begin{aligned} &t(\alpha_1\mathbf{v}_1 + \alpha_2\mathbf{v}_2 + \cdots + \alpha_n\mathbf{v}_n) \\ &\quad = \alpha_1 t(\mathbf{v}_1) + \alpha_2 t(\mathbf{v}_2) + \cdots + \alpha_n t(\mathbf{v}_n), \end{aligned}$$

for all $\mathbf{v}_1, \ldots, \mathbf{v}_n \in V$ and all $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$, and $n \in \mathbb{N}$.

7. The **zero transformation** and **identity transformation** are linear transformations:

$$\begin{aligned} t : V &\longrightarrow W \\ \mathbf{v} &\longmapsto \mathbf{0} \end{aligned} \quad \text{and} \quad \begin{aligned} i_V : V &\longrightarrow V \\ \mathbf{v} &\longmapsto \mathbf{v}. \end{aligned}$$

2 Matrices of linear transformations

8. Let V and W have dimensions n and m, and bases $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $F = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$. Let $t : V \longrightarrow W$ be a linear transformation, and let $\mathbf{A}$ be an $m \times n$ matrix such that

$$t(\mathbf{v})_F = \mathbf{A}\mathbf{v}_E, \quad \text{for each vector } \mathbf{v} \text{ in } V.$$

The **matrix representation** of t with respect to the bases E and F is $\mathbf{v}_E \longmapsto \mathbf{A}\mathbf{v}_E = t(\mathbf{v})_F$.

The **matrix** of t with respect to the bases E and F is $\mathbf{A}$.

Theorem C40

Let $t : V \longrightarrow W$ be a linear transformation, let $E = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for V and let $F = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ be a basis for W. Let

$$\begin{aligned} t(\mathbf{e}_1) &= (a_{11}, a_{21}, \ldots, a_{m1})_F, \\ t(\mathbf{e}_2) &= (a_{12}, a_{22}, \ldots, a_{m2})_F, \\ &\ \ \vdots \\ t(\mathbf{e}_n) &= (a_{1n}, a_{2n}, \ldots, a_{mn})_F. \end{aligned}$$

Then there is exactly one matrix of t with respect to the bases E and F, namely

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

9. Finding the matrix of a linear transformation t for bases E and F

Strategy C15 Matrix representation

To find the matrix $\mathbf{A}$ of a linear transformation $t : V \longrightarrow W$ with respect to the basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ for V, and the basis $F = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_m\}$ for W, do the following.

1. Find $t(\mathbf{e}_1), t(\mathbf{e}_2), \ldots, t(\mathbf{e}_n)$.
2. Find the F-coordinates of each of these image vectors.

$$\begin{aligned} t(\mathbf{e}_1) &= (a_{11}, a_{21}, \ldots, a_{m1})_F \\ t(\mathbf{e}_2) &= (a_{12}, a_{22}, \ldots, a_{m2})_F \\ &\vdots \\ t(\mathbf{e}_n) &= (a_{1n}, a_{2n}, \ldots, a_{mn})_F \end{aligned}$$

3. Construct the matrix $\mathbf{A}$ column by column.

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

The matrix representation is unique for given bases with vectors in a given order, however different matrix representations result from:

- different bases for V and W
- the basis vectors being in a different order.

10. Matrix definition of a linear transformation

Theorem C41

Let $t : V \longrightarrow W$ be a function that has a matrix representation. Then t is a linear transformation.

Corollary C42

A function $t : V \longrightarrow W$, where V and W are finite-dimensional vector spaces, is a linear transformation if and only if it has a matrix representation.

11. The linear transformations from $\mathbb{R}^2$ to itself are those functions that have a matrix representation

$$\begin{pmatrix} x \\ y \end{pmatrix} \longmapsto \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

for some $a, b, c, d \in \mathbb{R}$.

Similar expressions exist for linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^m$.

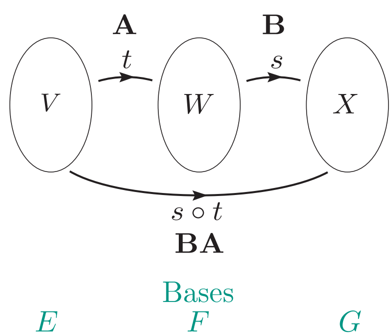
3 Composition and invertibility

12. Combining linear transformations

Theorem C43 Composition Rule

Let $t : V \longrightarrow W$ and $s : W \longrightarrow X$ be linear transformations. Then:

1. $s \circ t : V \longrightarrow X$ is a linear transformation
2. if $\mathbf{A}$ is the matrix of t with respect to the bases E and F, and $\mathbf{B}$ is the matrix of s with respect to the bases F and G, then $\mathbf{BA}$ is the matrix of $s \circ t$ with respect to the bases E and G.



Corollary C44

Let $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ be matrices of sizes $q \times p$, $p \times m$ and $m \times n$, respectively. Then

$$\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}.$$

13. An **invertible linear transformation** $t : V \longrightarrow W$ has an inverse function $t^{-1} : W \longrightarrow V$ such that

$$t^{-1} \circ t = i_V \quad \text{and} \quad t \circ t^{-1} = i_W.$$

A linear transformation is invertible if and only if it is one-to-one and onto.

Theorem C45 Inverse Rule

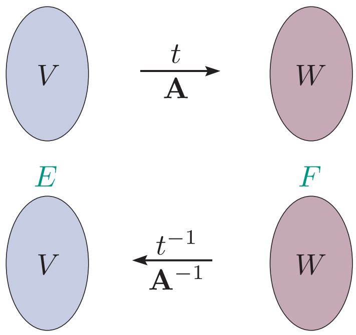
Let $t : V \longrightarrow W$ be a linear transformation.

(a) If t is invertible, then $t^{-1} : W \longrightarrow V$ is also a linear transformation.

(b) If $\mathbf{A}$ is the matrix of t with respect to the bases E and F, then:

(i) t is invertible if and only if $\mathbf{A}$ is invertible

(ii) if t is invertible, then $\mathbf{A}^{-1}$ is the matrix of t^{-1} with respect to the bases F and E.



Corollary C46

Let $t : V \longrightarrow W$ be an invertible linear transformation, where V and W are finite-dimensional. Then

$$\dim V = \dim W.$$

Strategy C16 Invertibility

To determine whether or not a linear transformation $t : V \longrightarrow W$ is invertible, where V and W are n-dimensional vector spaces with bases E and F, respectively, do the following.

1. Find a matrix representation of t,

$$\mathbf{v}_E \longmapsto \mathbf{A}\mathbf{v}_E = t(\mathbf{v})_F \quad \text{(Strategy C15).}$$

2. Evaluate $\det \mathbf{A}$ (Strategy C5).

- If $\det \mathbf{A} = 0$, then t is not invertible.
- If $\det \mathbf{A} \neq 0$, then t is invertible and $t^{-1} : W \longrightarrow V$ has the matrix representation

$$\mathbf{w}_F \longmapsto \mathbf{A}^{-1}\mathbf{w}_F = t^{-1}(\mathbf{w})_E.$$

14. Isomorphisms

The vector spaces V and W are **isomorphic** if there exists an invertible linear transformation $t : V \longrightarrow W$.

An **isomorphism** is such a function t.

Theorem C47

The finite-dimensional vector spaces V and W are isomorphic if and only if

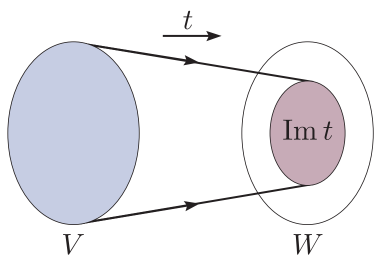
$$\dim V = \dim W.$$

In particular, each n-dimensional vector space is isomorphic to $\mathbb{R}^n$.

4 Image and kernel

15. The **image set** of a linear transformation $t : V \longrightarrow W$ is the set

$$\operatorname{Im} t = \{t(\mathbf{v}) : \mathbf{v} \in V\}.$$



Theorem C48

Let $t : V \longrightarrow W$ be a linear transformation. Then $\operatorname{Im} t$ is a subspace of the codomain W.

Proposition C49

A linear transformation t is onto if and only if $\dim(\operatorname{Im} t) = \dim W$.

16. Finding a basis for Im t

The images of the domain basis vectors span Im t.

Strategy C17 Basis for Im t

To find a basis for Im t, where $t : V \longrightarrow W$ is a linear transformation, do the following.

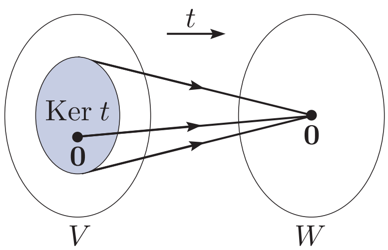
1. Find a basis $\{\mathbf{e}_1, \dots, \mathbf{e}_n\}$ for the domain V.
2. Determine the vectors $t(\mathbf{e}_1), \dots, t(\mathbf{e}_n)$.
3. If there is a vector $\mathbf{v}$ in $S = \{t(\mathbf{e}_1), \dots, t(\mathbf{e}_n)\}$ that is a linear combination of the other vectors in S, then discard $\mathbf{v}$ to give the set $S_1 = S - \{\mathbf{v}\}$.
4. If there is a vector $\mathbf{v}_1$ in S_1 such that $\mathbf{v}_1$ is a linear combination of the other vectors in S_1, then discard $\mathbf{v}_1$ to give the set $S_2 = S_1 - \{\mathbf{v}_1\}$.

Continue discarding vectors in this way until you obtain a linearly independent set. This set is a basis for Im t.

The dimension of Im t is equal to the number of vectors in the basis for Im t.

17. The **kernel** of a linear transformation $t : V \longrightarrow W$ is the set

$$\operatorname{Ker} t = \{\mathbf{v} \in V : t(\mathbf{v}) = \mathbf{0}\}.$$



Theorem C50

Let $t : V \longrightarrow W$ be a linear transformation. Then Ker t is a subspace of the domain V.

To find Ker t, solve the system of linear equations obtained by equating coordinates in $t(\mathbf{v}) = \mathbf{0}$.

Theorem C51 Solution Set Theorem

Let $t : V \longrightarrow W$ be a linear transformation. Let $\mathbf{b} \in W$ and let $\mathbf{a}$ be one vector in V that maps to $\mathbf{b}$, that is, $t(\mathbf{a}) = \mathbf{b}$. Then the solution set of the equation $t(\mathbf{x}) = \mathbf{b}$ is

$$\{\mathbf{x} : \mathbf{x} = \mathbf{a} + \mathbf{k} \text{ for some } \mathbf{k} \in \operatorname{Ker} t\}.$$

Proposition C52

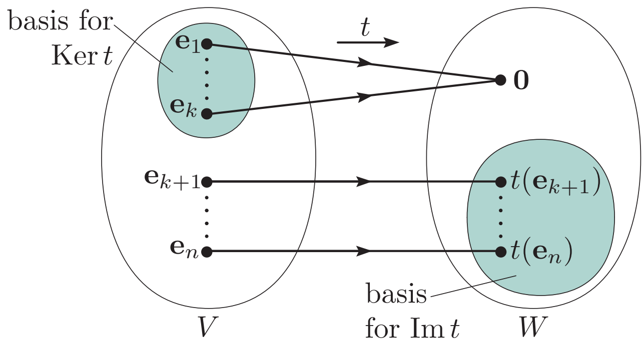
A linear transformation t is one-to-one if and only if Ker $t = \{\mathbf{0}\}$.

18. Dimension

Theorem C53 Dimension Theorem

Let $t : V \longrightarrow W$ be a linear transformation. Then

$$\dim(\operatorname{Im} t) + \dim(\operatorname{Ker} t) = \dim V.$$

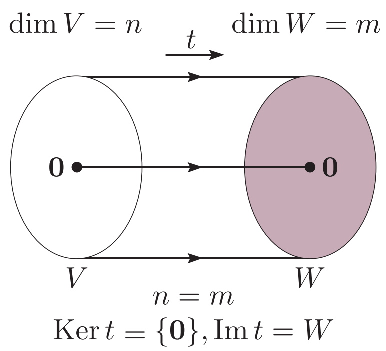


Theorem C54

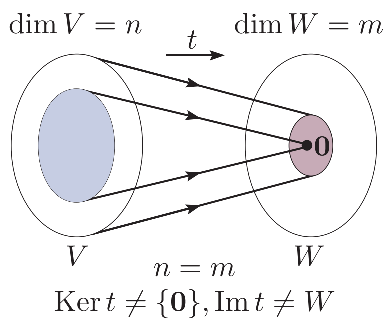
Let $t : V \longrightarrow W$ be a linear transformation from an n-dimensional vector space V to an m-dimensional vector space W.

1. If $n > m$, then t is not one-to-one: Ker $t \neq \{\mathbf{0}\}$.
2. If $n < m$, then t is not onto: Im $t \neq W$.
3. If $n = m$, then
   - *either* t is both one-to-one and onto: Ker $t = \{\mathbf{0}\}$ and Im $t = W$
   - *or* t is neither one-to-one nor onto: Ker $t \neq \{\mathbf{0}\}$ and Im $t \neq W$.

- The case where $\dim V = \dim W$ and t is both one-to-one and onto: $\operatorname{Ker} t = \mathbf{0}$ and $\operatorname{Im} t = W$.



- The case where $\dim V = \dim W$ and t is neither one-to-one nor onto: $\operatorname{Ker} t \neq \{\mathbf{0}\}$ and $\operatorname{Im} t \neq W$.



19. Number of solutions of a system of linear equations

> **Theorem C55**
>
> Let $\mathbf{Ax} = \mathbf{b}$ be a system of m linear equations in n unknowns.
>
> 1. If $n > m$, then $\mathbf{Ax} = \mathbf{b}$ has either no solution or infinitely many solutions.
> 2. If $n < m$, then there is some $\mathbf{b}$ for which $\mathbf{Ax} = \mathbf{b}$ has no solutions.
> 3. If $n = m$, then:
>   - *either* $\mathbf{Ax} = \mathbf{b}$ has exactly one solution for each $\mathbf{b}$
>   - *or* there are some $\mathbf{b}$ for which $\mathbf{Ax} = \mathbf{b}$ has no solutions; for all other $\mathbf{b}$, $\mathbf{Ax} = \mathbf{b}$ has infinitely many solutions.

Unit C4 Eigenvectors

Throughout, V denotes a vector space.

1 Eigenvectors and eigenvalues

1. Let $t : V \longrightarrow V$ be a linear transformation.

An **eigenvector** of t is a non-zero vector $\mathbf{v}$ that is mapped by t to a scalar multiple of itself:

$$t(\mathbf{v}) = \lambda\mathbf{v}, \quad \text{for some } \lambda \in \mathbb{R}.$$

The corresponding **eigenvalue** is this scalar λ.

We exclude the case $\mathbf{v} = \mathbf{0}$, since $t(\mathbf{0}) = \mathbf{0}$ for *every* linear transformation t. However, when $\lambda = 0$ the linear transformation maps every vector corresponding to this eigenvalue to the origin.

2. An **eigenvector** of a square matrix $\mathbf{A}$ is a non-zero vector $\mathbf{v}$ satisfying

$$\mathbf{Av} = \lambda\mathbf{v}, \quad \text{for some } \lambda \in \mathbb{R}.$$

The corresponding **eigenvalue** is λ.

The **eigenvector equations** are the equations $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$.

The **characteristic equation** of a square matrix $\mathbf{A}$ is the equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

The matrix $\mathbf{A} - \lambda\mathbf{I}$ is obtained by subtracting λ from each entry on the main diagonal of $\mathbf{A}$.

> **Strategy C18 Eigenvalues/vectors**
>
> To determine the eigenvalues and eigenvectors of a square matrix $\mathbf{A}$, do the following.
>
> 1. Find the eigenvalues:
>   - write down the characteristic equation
> $$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$
>   - expand this determinant to obtain a polynomial equation in λ (Strategy C5)
>   - solve this equation to find the eigenvalues.
> 2. Find the eigenvectors:
>   - write down the eigenvector equations
> $$(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$$
>   - for each eigenvalue λ, solve this system of linear equations to find the corresponding eigenvectors.

3. Diagonal entries of matrices

Proposition C56

The sum of the eigenvalues of a square matrix $\mathbf{A}$ is equal to the sum of the diagonal entries of $\mathbf{A}$ (the **trace** of $\mathbf{A}$).

Theorem C57

The eigenvalues of a triangular matrix and of a diagonal matrix are the diagonal entries of the matrix.

4. Eigenspaces

Theorem C58

Let $t : V \longrightarrow V$ be a linear transformation. For each eigenvalue λ of t, let $S(\lambda)$ be the set of vectors satisfying $t(\mathbf{v}) = \lambda\mathbf{v}$; that is, $S(\lambda)$ is the set of eigenvectors corresponding to λ, together with the zero vector $\mathbf{0}$. Then $S(\lambda)$ is a subspace of V.

$S(\lambda)$ is the **eigenspace** of t corresponding to the eigenvalue λ.

***Warning*:** $\mathbf{0}$ is in the eigenspace, but is not an eigenvector.

5. If the characteristic equation of a square matrix $\mathbf{A}$ can be written as

$$(\lambda - \lambda_1)^{m_1}(\lambda - \lambda_2)^{m_2} \cdots (\lambda - \lambda_p)^{m_p} = 0,$$

where $\lambda_1, \lambda_2, \ldots, \lambda_p$ are distinct, then the eigenvalue λ_j of $\mathbf{A}$ has **multiplicity** m_j, for $j = 1, 2, \ldots, p$.

The dimension of the eigenspace $S(\lambda_j)$ is at most equal to the multiplicity m_j of the corresponding eigenvalue λ_j, for $j = 1, 2, \ldots, p$. (However, the dimension can be less than the multiplicity m_j if m_j is greater than 1.)

2 Diagonalising matrices

6. Eigenvector bases

An **eigenvector basis** of a linear transformation $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a basis for $\mathbb{R}^n$ consisting of eigenvectors of t.

Strategy C19 Matrix representation V to V, same basis

To find the matrix $\mathbf{A}$ of a linear transformation $t : V \longrightarrow V$ with respect to the basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$:

1. find $t(\mathbf{e}_1), t(\mathbf{e}_2), \ldots, t(\mathbf{e}_n)$
2. find the E-coordinates of each of these image vectors
3. construct the matrix $\mathbf{A}$ column by column using the E-coordinates of $t(\mathbf{e}_j)$ to form column j, for $j = 1, 2, \ldots, n$.

Theorem C59

Let $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation, let $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be an eigenvector basis of t and let $t(\mathbf{e}_j) = \lambda_j\mathbf{e}_j$, for $j = 1, 2, \ldots, n$. Then the matrix of t with respect to the eigenvector basis E is

$$\mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

If the basis E used in Strategy C19 is an eigenvector basis, then the matrix of t is diagonal and can be written down using the corresponding eigenvalues.

7. Transition matrices

The **transition matrix P** from the basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ for $\mathbb{R}^n$ to the standard basis for $\mathbb{R}^n$ is the matrix whose jth column is formed from the standard coordinates of $\mathbf{e}_j$.

Theorem C60

Let $\mathbf{P}$ be the transition matrix from the basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ for $\mathbb{R}^n$ to the standard basis for $\mathbb{R}^n$. Then the standard coordinate representation of a vector in $\mathbb{R}^n$ is given by

$$\mathbf{v} = \mathbf{P}\mathbf{v}_E.$$

Moreover, $\mathbf{P}$ is invertible and

$$\mathbf{v}_E = \mathbf{P}^{-1}\mathbf{v}.$$

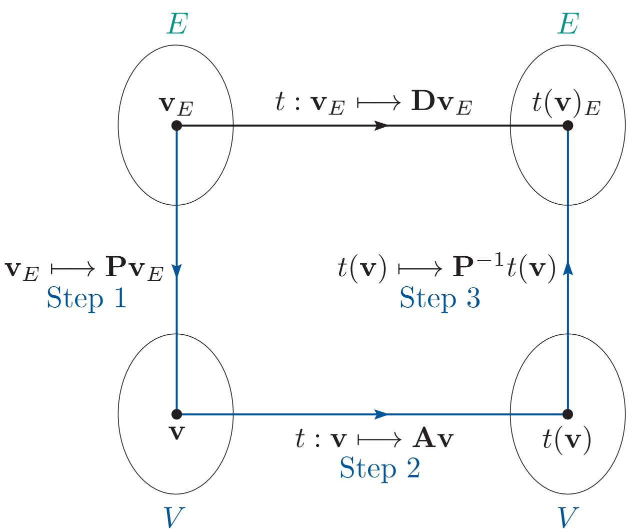
The transition matrix $\mathbf{P}$ is the matrix of the identity transformation i of $\mathbb{R}^n$ with respect to the basis E for the domain and the standard basis for the codomain.

When E is the standard basis for $\mathbb{R}^n$, the matrix $\mathbf{P}$ is the identity matrix $\mathbf{I}_n$.

Corollary C61

The rows or columns of an $n \times n$ matrix $\mathbf{A}$ form a set of n linearly independent vectors if and only if $\det \mathbf{A} \neq 0$.

8. Transition matrices link the matrix $\mathbf{A}$ of a linear transformation with respect to the standard basis, with the diagonal matrix $\mathbf{D}$ with respect to an eigenvector basis E.



$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P}\mathbf{v}_E = \mathbf{D}\mathbf{v}_E$$

Theorem C62

Let $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation and let E be an eigenvector basis of t. Let $\mathbf{A}$ be the matrix of t with respect to the standard basis for $\mathbb{R}^n$, let $\mathbf{D}$ be the matrix of t with respect to the eigenvector basis E and let $\mathbf{P}$ be the transition matrix from E to the standard basis for $\mathbb{R}^n$. Then

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}.$$

The matrices $\mathbf{D}$ and $\mathbf{A}$ are conjugate matrices.

9. The matrix $\mathbf{A}$ is **diagonalisable** if there exists an invertible matrix $\mathbf{P}$ such that the matrix

$$\mathbf{D} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

is diagonal.

10. An **eigenvector basis** of an $n \times n$ matrix $\mathbf{A}$ is a basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ for $\mathbb{R}^n$ consisting of eigenvectors of $\mathbf{A}$. Thus E is an eigenvector basis of the linear transformation $t : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ given by $t(\mathbf{v}) = \mathbf{A}\mathbf{v}$.

11. Diagonalising matrices

Strategy C20 Matrix diagonalisation

To diagonalise an $n \times n$ matrix $\mathbf{A}$:

1. find all the eigenvalues of $\mathbf{A}$ (Strategy C18)
2. find (if possible) an eigenvector basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbf{A}$ (Strategy C21)
3. write down the transition matrix $\mathbf{P}$ whose jth column is formed from the standard coordinates of $\mathbf{e}_j$.

Then

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_j is the eigenvalue corresponding to the eigenvector $\mathbf{e}_j$.

If $\mathbf{A}$ does not have an eigenvector basis, then $\mathbf{A}$ is not diagonalisable.

12. Finding an eigenvector basis

Theorem C63

Let $\mathbf{A}$ be an $n \times n$ matrix with *distinct* eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and corresponding eigenvectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. Then $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an eigenvector basis of $\mathbf{A}$.

To form an eigenvector basis take one eigenvector corresponding to each of the n distinct eigenvalues.

It may be possible to find an eigenvector basis of an $n \times n$ matrix even when it does not have n distinct eigenvalues.

Strategy C21 Eigenvector basis

To find an eigenvector basis of an $n \times n$ matrix $\mathbf{A}$:

1. find a basis for each eigenspace of $\mathbf{A}$
2. form the set E of all the basis vectors found in step 1.

If there are n vectors in E, then E is an eigenvector basis of $\mathbf{A}$; otherwise E is not a basis.

13. Calculating powers of matrices

Powers of a diagonalisable matrix $\mathbf{A}$ can be calculated using

$$\mathbf{A}^n = \mathbf{P}\mathbf{D}^n\mathbf{P}^{-1}, \quad \text{for } n = 1, 2, \ldots.$$

3 Symmetric matrices

14. An **orthonormal basis** for $\mathbb{R}^n$ consists of n mutually orthogonal vectors of magnitude 1.

An **orthogonal** matrix is an $n \times n$ matrix whose columns form an orthonormal basis for $\mathbb{R}^n$.

Warning: The columns of an orthogonal matrix are *orthonormal* vectors, not just *orthogonal* vectors.

15. A matrix $\mathbf{A}$ is **orthogonally diagonalisable** if there exists an orthogonal matrix $\mathbf{P}$ such that the matrix

$$\mathbf{D} = \mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{P}^{-1}\mathbf{A}\mathbf{P}$$

is diagonal.

16. A matrix $\mathbf{A}$ is orthogonally diagonalisable if and only if it is symmetric.

17. Diagonalising symmetric matrices

Strategy C22 Orthogonal diagonalisation

To orthogonally diagonalise an $n \times n$ symmetric matrix $\mathbf{A}$:

1. find all the eigenvalues of $\mathbf{A}$ (Strategy C18)
2. find an orthonormal eigenvector basis $E = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathbf{A}$ (Strategy C23)
3. write down the orthogonal transition matrix $\mathbf{P}$ whose jth column is formed from the standard coordinates of $\mathbf{e}_j$.

Then

$$\mathbf{P}^T\mathbf{A}\mathbf{P} = \mathbf{D} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix},$$

where λ_j is the eigenvalue corresponding to the eigenvector $\mathbf{e}_j$.

18. An orthonormal eigenvector basis

Theorem C64

Eigenvectors corresponding to *distinct* eigenvalues of a symmetric matrix are orthogonal.

The eigenspace of a symmetric matrix corresponding to an eigenvalue of multiplicity r has an orthonormal basis comprising r vectors.

Strategy C23 Orthonormal basis

To find an orthonormal eigenvector basis of a *symmetric* matrix $\mathbf{A}$:

1. find an orthonormal basis for each eigenspace of $\mathbf{A}$ (Strategy C13)
2. form the set E of all the basis vectors found in step 1.

Then E is an orthonormal eigenvector basis of $\mathbf{A}$.

19. Orthogonal matrices

Theorem C65

A square matrix $\mathbf{P}$ is orthogonal if and only if

$$\mathbf{P}^T = \mathbf{P}^{-1}.$$

Corollary C66

Let $\mathbf{P}$ and $\mathbf{Q}$ be orthogonal $n \times n$ matrices. Then:

- $\mathbf{P}^{-1}(= \mathbf{P}^T)$ is orthogonal
- the rows of $\mathbf{P}$ form an orthonormal basis for $\mathbb{R}^n$
- $\det \mathbf{P} = \pm 1$
- the product $\mathbf{PQ}$ is orthogonal.

20. Linear transformations of $\mathbb{R}^2$ whose matrices are orthogonal are rotations about the origin when the determinant is $+1$, and reflections in a line through the origin when the determinant is -1.

Linear transformations of $\mathbb{R}^3$ whose matrices are orthogonal are rotations about a line through the origin, reflections in a plane through the origin, or combinations of these. The orthogonal matrices representing rotations of $\mathbb{R}^3$ are precisely those with determinant $+1$.

21. Orthogonally diagonalising a linear transformation

Let t be a linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^n$ with a matrix representation that is a symmetric matrix $\mathbf{A}$. In effect, when we orthogonally diagonalise $\mathbf{A}$, we are finding a basis for $\mathbb{R}^n$ for which

- the matrix of t is diagonal
- the basis vectors are orthogonal
- the basis vectors have magnitude 1.

For $\mathbb{R}^2$ and $\mathbb{R}^3$ this new basis is simply the standard basis rotated, reflected or, for $\mathbb{R}^3$, a combination of the two.

4 Conics and quadrics

22. Conics

The three types of non-degenerate conic in standard position are illustrated in the Quick reference section on page 135.

The major and minor axes of a conic in standard position are the x-axis and y-axis, respectively.

The equation of a conic is in **standard form** if the major and minor axes of the conic align with the axes of the plane.

The equation of a conic in standard form resembles that of a conic in standard position.

The equations of the non-degenerate conics in standard position are:

- Ellipse $\quad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$
- Parabola $\quad y^2 = 4ax$
- Hyperbola $\quad \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1.$

Strategy C24 Classifying conics

To write the non-degenerate conic with equation

$$Ax^2 + Bxy + Cy^2 + Fx + Gy + H = 0$$

in standard form, do the following.

1. Introduce matrices:
   - write down
   $$\mathbf{A} = \begin{pmatrix} A & \frac{1}{2}B \\ \frac{1}{2}B & C \end{pmatrix} \quad \text{and} \quad \mathbf{J} = \begin{pmatrix} F \\ G \end{pmatrix}.$$
2. Align the axes:
   - orthogonally diagonalise $\mathbf{A}$ (Strategy C22) to get
   $$\mathbf{P}^T\mathbf{AP} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$
   - find $\begin{pmatrix} f & g \end{pmatrix} = \mathbf{J}^T\mathbf{P}$, and write the conic in the form
   $$\lambda_1(x')^2 + \lambda_2(y')^2 + fx' + gy' + H = 0.$$
3. Translate the origin:
   - complete the squares
   - make a substitution to change to the coordinate system (x'', y'').

23. Quadrics

The six types of quadric considered are illustrated (in standard position) in the Quick reference section on page 135.

A **quadric** in $\mathbb{R}^3$ is the set of points (x, y, z) that satisfy an equation of the form

$$\begin{aligned}&Ax^2 + By^2 + Cz^2 + Fxy + Gyz\\&\qquad + Hxz + Jx + Ky + Lz + M = 0,\end{aligned}$$

where A to M are real numbers, and A, B, C, F, G and H are not all 0.

There are nine types of quadric involving curved surfaces in $\mathbb{R}^3$.

A quadric is in **standard position** when its axes are aligned with the x-, y- and z-axes in a similar manner to the non-degenerate conics.

The equation of a quadric in **standard form** resembles that of a quadric in standard position.

The different types of quadric can be distinguished by the **curves of intersection** of the planes parallel to the coordinate planes that meet the quadric in a non-trivial intersection.

24. The curves of intersection of a **non-degenerate quadric** are non-degenerate conics.

There are five types of non-degenerate quadric and we also consider the elliptic cone.

The equations of the six types of quadric considered, in standard position, are:

- Ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$
- Elliptic paraboloid $z = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$
- Hyperbolic paraboloid $z = \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2}$
- Hyperboloid of one sheet $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$
- Hyperboloid of two sheets $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = -1$
- Elliptic cone $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0.$

Strategy C25 Classifying quadrics

To write the quadric with equation

$$\begin{aligned}&Ax^2 + By^2 + Cz^2 + Fxy + Gyz\\&\qquad + Hxz + Jx + Ky + Lz + M = 0\end{aligned}$$

in standard form, do the following.

1. Introduce matrices:
   - write down the matrices
   $$\mathbf{A} = \begin{pmatrix} A & \tfrac{1}{2}F & \tfrac{1}{2}H \\ \tfrac{1}{2}F & B & \tfrac{1}{2}G \\ \tfrac{1}{2}H & \tfrac{1}{2}G & C \end{pmatrix}$$
   $$\text{and}\quad \mathbf{J} = \begin{pmatrix} J \\ K \\ L \end{pmatrix}.$$
2. Align the axes:
   - orthogonally diagonalise $\mathbf{A}$ (Strategy C22) to get
   $$\mathbf{P}^T\mathbf{A}\mathbf{P} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$
   - find $\begin{pmatrix} f & g & h \end{pmatrix} = \mathbf{J}^T\mathbf{P}$, and write the quadric in the form
   $$\begin{aligned}&\lambda_1(x')^2 + \lambda_2(y')^2 + \lambda_3(z')^2\\&\qquad + fx' + gy' + hz' + M = 0.\end{aligned}$$
3. Translate the origin:
   - complete the squares
   - make a substitution to change to the coordinate system (x'', y'', z'').

Warning: It is the distribution of the plus and minus signs that distinguishes the types of quadric. For example, an equation of the form

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

represents a hyperboloid of one sheet, and an equation of the form

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$

represents a hyperboloid of two sheets; it is equivalent to the equation

$$-\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = -1.$$

Book D Analysis 1

Unit D1 Numbers

1 Real numbers

1. The set of **natural numbers**, **integers** and **rational numbers** (or **rationals**) are, respectively,

$$\mathbb{N} = \{1, 2, 3, \ldots\},$$
$$\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\},$$
$$\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{N}\}.$$

These numbers can all be represented on a number line.

2. The rationals have a natural **order** on the number line.

If a lies to the left of b on the number line, then we say that

a is **less than** b or b is **greater than** a

and we write these as **strict** inequalities

$$a < b \quad \text{or} \quad b > a.$$

We write the **weak** inequalities $a \leq b$ or $b \geq a$, if either $a < b$ or $a = b$.

3. A **decimal** is an expression of the form

$$\pm\, a_0.a_1a_2a_3\ldots = \pm\, a_0 + \frac{a_1}{10} + \frac{a_2}{10^2} + \frac{a_3}{10^3}\ldots,$$

where a_0 is a non-negative integer and a_n is a digit for each $n \in \mathbb{N}$. (A digit is in the set $\{0, 1, \ldots 9\}$.)

A **terminating** or **finite decimal** has only a finite number of the digits $a_1, a_2, \ldots$ that are non-zero.

A **non-terminating** or **infinite decimal** has infinitely many of the digits $a_1, a_2, \ldots$ that are non-zero.

A **recurring decimal** is a decimal with a repeating block of digits; for example, $0.863\,63\ldots = 0.8\overline{63}$. By definition $1 = 0.\overline{9} = 0.999\ldots$

4. Every rational number can be represented by a finite or a recurring decimal.

An **irrational** number is a real number which is not rational.

The set of irrational numbers consists of all the non-recurring decimals.

The set $\mathbb{R}$ of **real numbers** is the union of the set of rational numbers and the set of irrational numbers; it is the set of all terminating, recurring and non-recurring decimals.

We order two real numbers by examining their decimal representations and noticing the first place at which the digits differ.

The **real line** is the number line complete with both rational and irrational points.

There is a one-to-one correspondence between the points on the real line and the set $\mathbb{R}$ of real numbers. We often use the word 'point' to mean 'number' in this context.

Order properties of $\mathbb{R}$

Trichotomy Property If $a, b \in \mathbb{R}$, then *exactly one* of the following holds:

$$a < b \quad \text{or} \quad a = b \quad \text{or} \quad a > b.$$

Transitive Property If $a, b, c \in \mathbb{R}$, then

$$a < b \quad \text{and} \quad b < c \implies a < c.$$

Archimedean Property If $a \in \mathbb{R}$, then there is a positive integer n such that

$$n > a.$$

Density Property If $a, b \in \mathbb{R}$ and $a < b$, then there is a rational number x and an irrational number y such that

$$a < x < b \quad \text{and} \quad a < y < b.$$

5. Arithmetic in $\mathbb{R}$

The set $\mathbb{R}$ of real numbers forms a **field**: the properties A1–A5, M1–M5 and D1 all hold. (See **4.** on page 14.)

More succinctly these properties mean that:

- $\mathbb{R}$ is an abelian group under the operation of addition $+$
- $\mathbb{R}^* = \mathbb{R} - \{0\}$ is an abelian group under the operation of multiplication $\times$;

and these two group structures are linked by the distributive property.

2 Inequalities

6. Rearranging inequalities

We use the usual rules for the sign of a product.

$\times$	$+$	$-$
$+$	$+$	$-$
$-$	$-$	$+$

In particular, the square of any real number is non-negative.

Rules for rearranging inequalities

Let a, b, c and p be real numbers.

Rule 1 $a < b \iff b - a > 0$.

Rule 2 $a < b \iff a + c < b + c$.

Rule 3 If $c > 0$, then $a < b \iff ac < bc$;
if $c < 0$, then $a < b \iff ac > bc$.

Rule 4 If $a, b > 0$, then

$$a < b \iff \frac{1}{a} > \frac{1}{b}.$$

Rule 5 If $a, b \geq 0$ and $p > 0$, then

$$a < b \iff a^p < b^p.$$

Rule 6 $|a| < b \iff -b < a < b$.

There are corresponding versions of these rules where the strict inequality $a < b$ or $|a| < b$ is replaced by the weak inequality $a \leq b$ or $|a| \leq b$, respectively.

7. The **solution set** of an inequality involving an unknown real number x is the set of values of x for which the given inequality holds.

To **solve** the inequality, find the solution set by rewriting the inequality in an equivalent but simpler form, using the rules for rearranging inequalities.

8. The **modulus** (or **absolute value**) of a real number a is defined by

$$|a| = \begin{cases} a, & \text{if } a \geq 0, \\ -a, & \text{if } a < 0. \end{cases}$$

The **distance** on the real line from 0 to a is $|a|$, and from a to b is $|a - b|$ $(= |b - a|)$.

Properties of the modulus

If $a, b \in \mathbb{R}$, then

1. $|a| \geq 0$, with equality if and only if $a = 0$
2. $-|a| \leq a \leq |a|$
3. $|a|^2 = a^2$
4. $|a - b| = |b - a|$
5. $|ab| = |a||b|$.

3 Proving inequalities

9. Rules for inequalities

Combination Rules

If $a < b$ and $c < d$, then

Sum Rule $a + c < b + d$

Product Rule $ac < bd$, provided $a, c \geq 0$.

Transitive Rule

$a < b$ and $b < c \implies a < c$.

Triangle Inequality

If $a, b \in \mathbb{R}$, then

1. $|a + b| \leq |a| + |b|$ (usual form)
2. $|a - b| \geq \big||a| - |b|\big|$ ('backwards' form).

There are also versions of the Combination and Transitive Rules involving weak inequalities.

There is a version of the Triangle Inequality for n real numbers, where $n > 2$:

$$|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n|.$$

10. Inequalities for real numbers

- $ab \leq \left(\dfrac{a+b}{2}\right)^2$, for $a, b \in \mathbb{R}$.
- $\sqrt{a^2 + b^2} \leq a + b$, for $a, b \geq 0$.
- $|\sqrt{a} - \sqrt{b}| \leq \sqrt{|a - b|}$, for $a, b \geq 0$.

11. The **binomial coefficient**

$$\binom{n}{k} = \frac{n!}{k!\,(n-k)!},$$

for any non-negative integers n and k with $k \leq n$, is the number of combinations of n objects taken k at a time. (Note $0! = 1$.)

Theorem D1 Binomial Theorem

If $a, b \in \mathbb{R}$ and $n \in \mathbb{N}$, then

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$
$$= a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \cdots + b^n.$$

In the important special case where $a = 1$ and $b = x \in \mathbb{R}$, we have

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$$
$$= 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots + x^n.$$

12. Inequalities for natural numbers

Theorem D2 Bernoulli's Inequality

If $x \in \mathbb{R}$ and $n \in \mathbb{N}$, then

$$(1+x)^n \geq 1 + nx, \quad \text{when } x \geq -1.$$

- $2^n \geq 1 + n$, for $n \geq 1$.
- $2^n \geq n^2$, for $n \geq 4$.
- $1 + \dfrac{1}{2n-1} \leq 2^{1/n} \leq 1 + \dfrac{1}{n}$, for $n \geq 1$.

4 Least upper bounds

13. A set $E \subseteq \mathbb{R}$ is **bounded above** if there is a real number M, called an **upper bound** of E, such that

$$x \leq M, \quad \text{for all } x \in E.$$

If the upper bound M belongs to E, then M is called the **maximum element** of E, and is denoted by $\max E$.

14. A set $E \subseteq \mathbb{R}$ is **bounded below** if there is a real number m, called a **lower bound** of E, such that

$$m \leq x, \quad \text{for all } x \in E.$$

If the lower bound m belongs to E, then m is called the **minimum element** of E, and is denoted by $\min E$.

15. A set $E \subseteq \mathbb{R}$ is **bounded** if E is bounded above and bounded below; the set E is **unbounded** if it is not bounded.

16. A real number M is the **least upper bound**, or **supremum**, of a set $E \subseteq \mathbb{R}$ if

1. M is an upper bound of E
2. each number $M' < M$ is not an upper bound of E.

In this case, we write $M = \sup E$.

Strategy D1 Least upper bound

To show that M is the least upper bound (supremum) of a subset E of $\mathbb{R}$, check that:

1. $x \leq M$, for *all* $x \in E$
2. if $M' < M$, then there is *some* $x \in E$ such that $x > M'$.

Least Upper Bound Property of $\mathbb{R}$

Let E be a non-empty subset of $\mathbb{R}$. If E is bounded above, then E has a least upper bound.

17. A real number m is the **greatest lower bound**, or **infimum**, of a set $E \subseteq \mathbb{R}$ if

1. m is a lower bound of E
2. each number $m' > m$ is not a lower bound of E.

In this case, we write $m = \inf E$.

Strategy D2 Greatest lower bound

To show that m is the greatest lower bound (infimum) of a subset E of $\mathbb{R}$, check that:

1. $x \geq m$, for *all* $x \in E$
2. if $m' > m$, then there is *some* $x \in E$ such that $x < m'$.

Greatest Lower Bound Property of $\mathbb{R}$

Let E be a non-empty subset of $\mathbb{R}$. If E is bounded below, then E has a greatest lower bound.

5 Manipulating real numbers

18. We define the **sum** and **product** of two positive real numbers a and b (expressed as decimals) as follows. Form the sums (or products) of truncations of a and b to n decimal places for each $n \in \mathbb{N}$, and take the least upper bound of the resulting set of finite decimals.

Similar ideas can be used to define the operations of subtraction and division.

19. Existence of roots

> **Theorem D3**
>
> For each positive real number a and each integer $n > 1$, there is a unique positive real number b such that
>
> $$b^n = a.$$

We call this positive number b the ***n*th root** of a, and write $b = \sqrt[n]{a}$. We also define $\sqrt[n]{0} = 0$, since $0^n = 0$.

For each positive number a there are two real solutions to the equation $b^n = a$ when n is even: $\sqrt[n]{a}$ and $a^{1/n}$ always denote the *positive* nth root of a. We write $\pm\sqrt[n]{a}$ or $\pm a^{1/n}$ to refer to both real solutions (for example, when solving equations).

20. If $a > 0$, $m \in \mathbb{Z}$ and $n \in \mathbb{N}$, then

$$a^{m/n} = (\sqrt[n]{a})^m = \sqrt[n]{a^m}.$$

> **Index Laws**
>
> If $a, b > 0$ and $x, y \in \mathbb{Q}$, then
>
> - $a^x b^x = (ab)^x$
> - $a^x a^y = a^{x+y}$
> - $(a^x)^y = a^{xy}$.

Unit D2 Sequences

1 Introducing sequences

1. A **sequence** is an unending list of real numbers

$$a_1, a_2, a_3, \ldots .$$

The real number a_n is called the ***n*th term** of the sequence, and the sequence is denoted by

$$(a_n).$$

A sequence may begin with a term other than a_1; for example, a_0 or a_3.

2. A **sequence diagram** is the graph of the function from $\mathbb{N}$ to $\mathbb{R}$ that represents the sequence; that is, the following set of points in $\mathbb{R}^2$:

$$\{(n, a_n) : n = 1, 2, \ldots\}.$$

3. A sequence (a_n) is said to be

- **constant** if
  $$a_{n+1} = a_n, \quad \text{for } n = 1, 2, \ldots$$
- **increasing** if
  $$a_{n+1} \geq a_n, \quad \text{for } n = 1, 2, \ldots$$
- **strictly increasing** if
  $$a_{n+1} > a_n, \quad \text{for } n = 1, 2, \ldots$$
- **decreasing** if
  $$a_{n+1} \leq a_n, \quad \text{for } n = 1, 2, \ldots$$
- **strictly decreasing** if
  $$a_{n+1} < a_n, \quad \text{for } n = 1, 2, \ldots$$
- **monotonic** if (a_n) is either increasing or decreasing
- **strictly monotonic** if (a_n) is either strictly increasing or strictly decreasing.

> **Strategy D3 Monotonic sequence**
>
> To show that a given sequence (a_n) is monotonic, consider the difference $a_{n+1} - a_n$.
>
> - If $a_{n+1} - a_n \geq 0$, for $n = 1, 2, \ldots$, then (a_n) is increasing.
> - If $a_{n+1} - a_n \leq 0$, for $n = 1, 2, \ldots$, then (a_n) is decreasing.

Strategy D4 Monotonic (+ve terms)

To show that a given sequence (a_n) of *positive* terms is monotonic, consider the quotient $\dfrac{a_{n+1}}{a_n}$.

- If $\dfrac{a_{n+1}}{a_n} \geq 1$, for $n = 1, 2, \ldots$, then (a_n) is increasing.
- If $\dfrac{a_{n+1}}{a_n} \leq 1$, for $n = 1, 2, \ldots$, then (a_n) is decreasing.

4. A sequence (a_n) has a certain property **eventually** if the sequence has this property provided we ignore a finite number of terms.

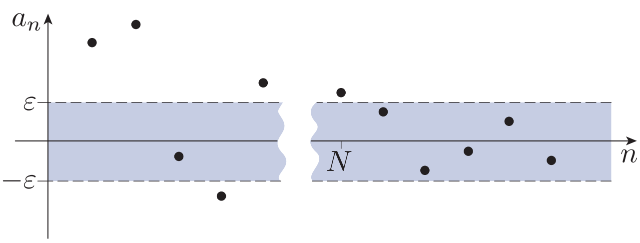
2 Null sequences

5. Sequences converging to 0

The sequence (a_n) is **null** if

> for each positive number ε, there is an integer N such that
>
> $|a_n| < \varepsilon, \quad \text{for all } n > N.$

In this case we also say that the sequence (a_n) is **convergent with limit 0**, or that (a_n) **converges** to 0.



Strategy D5 Null/non-null sequence

- To show that the sequence (a_n) is null, rearrange the inequality $|a_n| < \varepsilon$ to find an integer N (generally depending on ε) such that $|a_n| < \varepsilon$, for all $n > N$.
- To show that the sequence (a_n) is not null, find *one* value of $\varepsilon > 0$ for which there is *no* integer N such that $|a_n| < \varepsilon$ for all $n > N$.

6. The sequence (a_n) is null if and only if the sequence $(|a_n|)$ is null.

A null sequence (a_n) remains null if a finite number of terms are added, deleted or altered.

7. A sequence (a_n) is **dominated** by a sequence (b_n) of *non-negative* terms if

$$|a_n| < b_n, \quad \text{for } n = 1, 2, \ldots.$$

8. Rules for null sequences

Theorem D4 Power Rule

If (a_n) is null, where $a_n \geq 0$, for $n = 1, 2, \ldots$, and p is a positive real number, then (a_n^p) is null.

Theorem D5 Combination Rules

If (a_n) and (b_n) are null, then:

Sum Rule $\quad (a_n + b_n)$ is null

Multiple Rule $\quad (\lambda a_n)$ is null, $\quad$ for $\lambda \in \mathbb{R}$

Product Rule $\quad (a_n b_n)$ is null.

Theorem D6 Squeeze Rule

If (b_n) is a null sequence of non-negative terms, and

$$|a_n| \leq b_n, \quad \text{for } n = 1, 2, \ldots,$$

then (a_n) is null.

Strategy D6 Null sequence - Squeeze Rule

To use the Squeeze Rule to show that a sequence (a_n) is null, do the following.

1. Guess a dominating null sequence (b_n) with non-negative terms.
2. Check that $|a_n| \leq b_n$, for $n = 1, 2, \ldots$.

9. **Basic null sequences**

Theorem D7 Basic null sequences

The following sequences are null.

- $(1/n^p)$, for $p > 0$.
- (c^n), for $|c| < 1$.
- $(n^p c^n)$, for $p > 0, |c| < 1$.
- $(c^n/n!)$, for $c \in \mathbb{R}$.
- $(n^p/n!)$, for $p > 0$.

3 Convergent sequences

10. The sequence (a_n) is **convergent** with **limit l** if $(a_n - l)$ is a null sequence.

We say that (a_n) **converges to l** and we write

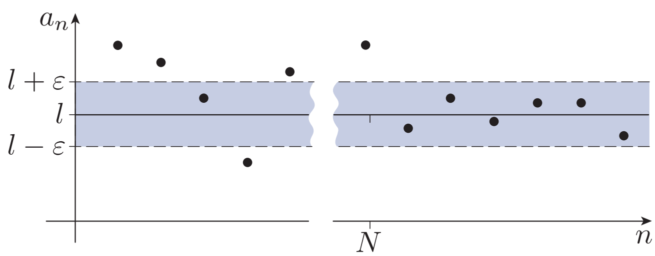
$$\lim_{n\to\infty} a_n = l \quad \text{or} \quad a_n \to l \ \text{ as } n \to \infty,$$

or simply $a_n \to l$.

Equivalently, the sequence (a_n) **converges to l** if

for each positive number ε, there is an integer N such that

$$|a_n - l| < \varepsilon, \quad \text{for all } n > N.$$



11. If a sequence is convergent, then it has a unique limit (See Corollary D12, in **15.** on page 81.)

A convergent sequence (a_n) with limit l remains convergent with limit l if a finite number of terms are added, deleted or altered.

12. **Sequences converging to 1**

- $\lim_{n\to\infty} n^{1/n} = 1$
- $\lim_{n\to\infty} a^{1/n} = 1$, for any positive number a.

13. The **dominant term** of a quotient involving the variable n, where $n = 1, 2, \ldots$, is the term in n (without its coefficient) which eventually has the largest absolute value.

Strategy D7 Limit of a sequence involving a quotient

To evaluate the limit of a sequence whose nth term is a complicated quotient, do the following.

1. Identify the dominant term, noting that

   $n!$ dominates c^n,

   and, for $|c| > 1$ and $p > 0$,

   c^n dominates n^p.
2. Divide both the numerator and denominator by the dominant term.
3. Apply the Combination Rules.

14. **Rules for convergent sequences**

Theorem D8 Combination Rules

If $\lim_{n\to\infty} a_n = l$ and $\lim_{n\to\infty} b_n = m$, then:

Sum Rule $\lim_{n\to\infty} (a_n + b_n) = l + m$

Multiple Rule $\lim_{n\to\infty} (\lambda a_n) = \lambda l$, for $\lambda \in \mathbb{R}$

Product Rule $\lim_{n\to\infty} (a_n b_n) = lm$

Quotient Rule $\lim_{n\to\infty} \left(\dfrac{a_n}{b_n}\right) = \dfrac{l}{m}$, provided that $m \neq 0$.

In applications of the Quotient Rule, (b_n) is always eventually non-zero since $m \neq 0$, although some terms b_n may be zero.

Lemma D9

If $\lim_{n\to\infty} a_n = l$ and $l > 0$, then there is an integer N such that

$$a_n > \tfrac{1}{2} l, \quad \text{for all } n > N.$$

15. Further rules for convergent sequences

Theorem D10 Squeeze Rule

If (a_n), (b_n) and (c_n) are sequences such that

1. $b_n \leq a_n \leq c_n$, for $n = 1, 2, \ldots,$
2. $\lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = l,$

then $\lim_{n\to\infty} a_n = l$.

Theorem D11 Limit Inequality Rule

If $\lim_{n\to\infty} a_n = l$ and $\lim_{n\to\infty} b_n = m$, and also

$$a_n \leq b_n, \quad \text{for } n = 1, 2, \ldots,$$

then $l \leq m$.

Note, however, that strict inequalities are not necessarily preserved in the Limit Inequality Rule.

Corollary D12

If $\lim_{n\to\infty} a_n = l$ and $\lim_{n\to\infty} a_n = m$, then $l = m$.

Theorem D13

If $\lim_{n\to\infty} a_n = l$, then $\lim_{n\to\infty} |a_n| = |l|$.

4 Divergent sequences

16. A sequence is **divergent** if it is not convergent.

A sequence (a_n) is **bounded** if there is a number M such that

$$|a_n| \leq M, \quad \text{for } n = 1, 2, \ldots.$$

A sequence is **unbounded** if it is not bounded.

Theorem D14

If (a_n) is convergent, then (a_n) is bounded.

Corollary D15

If (a_n) is unbounded, then (a_n) is divergent.

17. The sequence (a_n) **tends to infinity** if

for each positive number M, there is an integer N such that

$$a_n > M, \quad \text{for all } n > N.$$

In this case, we write

$$a_n \to \infty \text{ as } n \to \infty \quad \text{or simply} \quad a_n \to \infty.$$

The sequence (a_n) **tends to minus infinity** if

$$-a_n \to \infty \text{ as } n \to \infty.$$

We write

$$a_n \to -\infty \text{ as } n \to \infty \quad \text{or simply} \quad a_n \to -\infty.$$

18. If a sequence tends to infinity, or to minus infinity, then it is unbounded and hence divergent.

A sequence (a_n) that tends to infinity remains a sequence that tends to infinity if a finite number of terms are added, deleted or altered. Similarly for a sequence that tends to minus infinity.

19. Rules for sequences tending to infinity

Theorem D16 Reciprocal Rule

If the sequence (a_n) satisfies the conditions

1. (a_n) is eventually positive
2. $(1/a_n)$ is a null sequence

then $a_n \to \infty$.

Theorem D17 Combination Rules

If (a_n) tends to infinity and (b_n) tends to infinity, then

Sum Rule	$(a_n + b_n)$ tends to infinity
Multiple Rule	(λa_n) tends to infinity for $\lambda \in \mathbb{R}^+$
Product Rule	$(a_n b_n)$ tends to infinity.

Theorem D18 Squeeze Rule

If (b_n) tends to infinity and

$$a_n \geq b_n, \quad \text{for } n = 1, 2, \ldots,$$

then (a_n) tends to infinity.

20. A **subsequence** of the sequence (a_n) is a sequence (a_{n_k}) where (n_k) is a strictly increasing sequence of positive integers; that is,

$$n_1 < n_2 < n_3 < \cdots.$$

- The **even subsequence** (a_{2k}) comprises the even terms of (a_n); that is, $a_2, a_4, \ldots, a_{2k}, \ldots$.
- The **odd subsequence** (a_{2k-1}) comprises the odd terms of (a_n); that is, $a_1, a_3, \ldots, a_{2k-1}, \ldots$.

A sequence is always a subsequence of itself.

Theorem D19

For any subsequence (a_{n_k}) of a sequence (a_n):

1. if $a_n \to l$ as $n \to \infty$, then $a_{n_k} \to l$ as $k \to \infty$
2. if $a_n \to \infty$ as $n \to \infty$, then $a_{n_k} \to \infty$ as $k \to \infty$.

Corollary D20 Subsequence Rules

First Subsequence Rule The sequence (a_n) is divergent if (a_n) has two convergent subsequences with different limits.

Second Subsequence Rule The sequence (a_n) is divergent if (a_n) has a subsequence which tends to infinity or a subsequence which tends to minus infinity.

Strategy D8 Divergent sequences

To prove that the sequence (a_n) is divergent, do one of the following:

- show that (a_n) has two convergent subsequences with different limits
- show that (a_n) has a subsequence which tends to infinity or a subsequence which tends to minus infinity.

21. A sequence (a_n) consists of two subsequences (a_{m_k}) and (a_{n_k}) if every term of the sequence appears in one or other of the subsequences.

Every sequence (a_n) consists of its even subsequence (a_{2k}) and its odd subsequence (a_{2k-1}).

Theorem D21

Let (a_n) consist of two subsequences (a_{m_k}) and (a_{n_k}), which both tend to the *same* limit l. Then

$$\lim_{n\to\infty} a_n = l.$$

5 Monotone Convergence Theorem

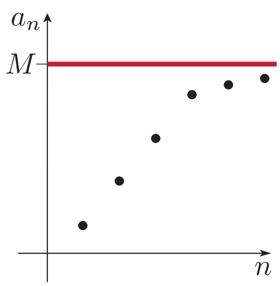
22. Monotonic sequences

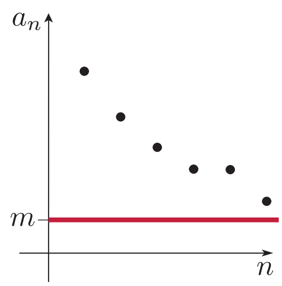
Theorem D22 Monotone Convergence Theorem

If the sequence (a_n) is either

- increasing and bounded above, or
- decreasing and bounded below,

then (a_n) is convergent.





Theorem D23 Monotonic Sequence Theorem

If the sequence (a_n) is monotonic, then either (a_n) is convergent or $a_n \to \pm\infty$.

23. The irrational numbers π and e

$$\pi = \lim_{n\to\infty} \tfrac{1}{2} s_n \sin(2\pi/s_n),$$

where $s_n = 3 \times 2^n$.

$$e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$$

and

$$e^x = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n, \quad \text{for any } x > 0.$$

Unit D3 Series

1 Introducing series

1. An **infinite series**, or **series**, is an expression

$$a_1 + a_2 + a_3 + \cdots$$

where (a_n) is a sequence.

The ***n*th term** of the series is a_n.

The ***n*th partial sum** of the series is

$$s_n = a_1 + a_2 + \cdots + a_n.$$

2. A series is **convergent** with **sum** s (or **converges to** s) if its sequence (s_n) of partial sums converges to s: we write

$$a_1 + a_2 + a_3 + \cdots = s.$$

A series **diverges**, or is **divergent**, if the sequence (s_n) diverges.

3. Sigma notation

We write

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots .$$

When using sigma notation to represent the nth partial sum of such a series, we write

$$s_n = a_1 + a_2 + \cdots + a_n = \sum_{k=1}^{n} a_k.$$

If a series begins with a term other than a_1, we write, for example,

$$\sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + \cdots ;$$

for such a series,

$$s_n = a_0 + a_1 + \cdots + a_n = \sum_{k=0}^{n} a_k.$$

4. Geometric series

Sum of a finite geometric series

The geometric series with first term a, common ratio $r \neq 1$ and n terms has the sum

$$a + ar + \cdots + ar^{n-1} = \frac{a(1-r^n)}{1-r}.$$

The **(infinite) geometric series** with first term a and common ratio r is

$$\sum_{n=0}^{\infty} ar^n = a + ar + ar^2 + \cdots .$$

Theorem D24 Geometric series

(a) If $|r| < 1$, then $\sum_{n=0}^{\infty} ar^n$ is convergent, with sum $\frac{a}{1-r}$.

(b) If $|r| \geq 1$ and $a \neq 0$, then $\sum_{n=0}^{\infty} ar^n$ is divergent.

5. A **telescoping series** is a series where nearly all of the terms in the partial sums cancel out.

6. Rules for convergent series

Theorem D25 Combination Rules

Suppose that $\sum_{n=1}^{\infty} a_n = s$ and $\sum_{n=1}^{\infty} b_n = t$. Then

Sum Rule $\quad \sum_{n=1}^{\infty} (a_n + b_n) = s + t$

Multiple Rule $\quad \sum_{n=1}^{\infty} \lambda a_n = \lambda s, \quad$ for $\lambda \in \mathbb{R}$.

7. Rule for divergent series

Corollary D26 Multiple Rule

Suppose that $\sum_{n=1}^{\infty} a_n$ is divergent and that λ is a non-zero real number.

Then $\sum_{n=1}^{\infty} \lambda a_n$ is divergent.

8. Non-null Test

Theorem D27

If $\sum_{n=1}^{\infty} a_n$ is a convergent series, then its sequence of terms (a_n) is a null sequence.

Corollary D28 Non-null Test

If (a_n) is not a null sequence, then $\sum_{n=1}^{\infty} a_n$ is divergent.

Strategy D9 Divergent - Non-null Test

To show that $\sum_{n=1}^{\infty} a_n$ is divergent using the Non-null Test, check that the sequence (a_n) is not null by showing that $(|a_n|)$ has one of the following:

- a convergent subsequence with non-zero limit
- a subsequence which tends to infinity.

Warning: The Non-null Test cannot be used to show that a series is convergent.

2 Series with non-negative terms

9. Comparison Test for series

Theorem D30 Comparison Test

(a) If $0 \leq a_n \leq b_n$, for $n = 1, 2, \ldots$, and $\sum_{n=1}^{\infty} b_n$ is convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent.

(b) If $0 \leq b_n \leq a_n$, for $n = 1, 2, \ldots$, and $\sum_{n=1}^{\infty} b_n$ is divergent, then $\sum_{n=1}^{\infty} a_n$ is divergent.

Strategy D10 Convergent/divergent - Comparison Test

- To show that a series $\sum_{n=1}^{\infty} a_n$ of non-negative terms is *convergent* using the Comparison Test, do the following.
  1. Guess that $\sum_{n=1}^{\infty} a_n$ is dominated by a convergent series $\sum_{n=1}^{\infty} b_n$.
  2. Check that $0 \leq a_n \leq b_n$, for $n = 1, 2, \ldots$.
- To show that a series $\sum_{n=1}^{\infty} a_n$ of non-negative terms is *divergent* using the Comparison Test, do the following.
  1. Guess that $\sum_{n=1}^{\infty} a_n$ dominates a divergent series $\sum_{n=1}^{\infty} b_n$.
  2. Check that $0 \leq b_n \leq a_n$, for $n = 1, 2, \ldots$.

The inequalities can be shown to hold *eventually*; that is, for all $n > N$, for some number N.

10. Limit Comparison Test for series

Theorem D31 Limit Comparison Test

Suppose that $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ have positive terms and that

$$\frac{a_n}{b_n} \to L \quad \text{as } n \to \infty,$$

where $L \neq 0$.

(a) If $\sum_{n=1}^{\infty} b_n$ is convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent.

(b) If $\sum_{n=1}^{\infty} b_n$ is divergent, then $\sum_{n=1}^{\infty} a_n$ is divergent.

Strategy D11 Convergent/divergent - Limit Comparison Test

- To show that a series $\sum_{n=1}^{\infty} a_n$ of positive terms is *convergent* using the Limit Comparison Test, do the following.
  1. Guess that $\sum_{n=1}^{\infty} a_n$ behaves like a comparable convergent series $\sum_{n=1}^{\infty} b_n$ of positive terms.
  2. Check that
  $$\frac{a_n}{b_n} \to L \neq 0 \text{ as } n \to \infty$$
  and deduce that $\sum_{n=1}^{\infty} a_n$ converges.
- To show that a series $\sum_{n=1}^{\infty} a_n$ of positive terms is *divergent* using the Limit Comparison Test, do the following.
  1. Guess that $\sum_{n=1}^{\infty} a_n$ behaves like a comparable divergent series $\sum_{n=1}^{\infty} b_n$ of positive terms.
  2. Check that
  $$\frac{a_n}{b_n} \to L \neq 0 \text{ as } n \to \infty$$
  and deduce that $\sum_{n=1}^{\infty} a_n$ diverges.

11. Ratio Test for series

Theorem D32 Ratio Test

Suppose that $\sum_{n=1}^{\infty} a_n$ has positive terms and that $\dfrac{a_{n+1}}{a_n} \to l$ as $n \to \infty$.

(a) If $0 \leq l < 1$, then $\sum_{n=1}^{\infty} a_n$ is convergent.

(b) If $l > 1$ or $l = \infty$, then $\sum_{n=1}^{\infty} a_n$ is divergent.

If $l = 1$, then the Ratio Test is inconclusive – the series may be convergent or divergent.

12. Basic Series

Theorem D33 Basic series

The following series are convergent:

- $\sum_{n=1}^{\infty} \dfrac{1}{n^p}$, for $p \geq 2$
- $\sum_{n=1}^{\infty} c^n$, for $0 \leq c < 1$
- $\sum_{n=1}^{\infty} n^p c^n$, for $p > 0$, $0 \leq c < 1$
- $\sum_{n=1}^{\infty} \dfrac{c^n}{n!}$, for $c \geq 0$.

The following series is divergent:

- $\sum_{n=1}^{\infty} \dfrac{1}{n^p}$, for $0 < p \leq 1$.

In particular, the **harmonic series**

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

is divergent.

3 Series with positive and negative terms

13. The series $\sum_{n=1}^{\infty} a_n$ is **absolutely convergent** if $\sum_{n=1}^{\infty} |a_n|$ is convergent.

Theorem D34 Absolute Convergence Test

If $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then $\sum_{n=1}^{\infty} a_n$ is convergent.

Theorem D35 Triangle Inequality (infinite form)

If $\sum_{n=1}^{\infty} a_n$ is absolutely convergent, then

$$\left|\sum_{n=1}^{\infty} a_n\right| \leq \sum_{n=1}^{\infty} |a_n|.$$

14. Alternating Test for series

Theorem D36 Alternating Test

Let

$$a_n = (-1)^{n+1} b_n, \quad n = 1, 2, \ldots,$$

where (b_n) is a decreasing null sequence with positive terms. Then

$$\sum_{n=1}^{\infty} a_n = b_1 - b_2 + b_3 - b_4 + \cdots$$

is convergent.

Strategy D12 Convergent series - Alternating Test

To prove that $\sum_{n=1}^{\infty} a_n$ is convergent using the Alternating Test, check that

$$a_n = (-1)^{n+1} b_n, \quad n = 1, 2, \ldots,$$

where

1. $b_n \geq 0$, for $n = 1, 2, \ldots$
2. (b_n) is a null sequence
3. (b_n) is decreasing.

15. Determining convergence/divergence

Strategy D13 General strategy

To determine whether a series $\sum a_n$ is convergent or divergent, do the following.

1. If you think that the sequence of terms (a_n) is non-null, then try the **Non-null Test.**
2. If $\sum a_n$ (eventually) only has non-negative terms, then try one of these tests.

   Basic series Is $\sum a_n$ a basic series, or a combination of these?

   Comparison Test Is $a_n \leq b_n$, where $\sum b_n$ is convergent, or $a_n \geq b_n \geq 0$, where $\sum b_n$ is divergent?

   Limit Comparison Test Does a_n behave like b_n for large n (that is, does $a_n/b_n \to L \neq 0$), where $\sum b_n$ is a series that you know converges or diverges?

   Ratio Test Does $a_{n+1}/a_n \to l \neq 1$?
3. If $\sum a_n$ has infinitely many positive and negative terms, then try one of these tests.

   Absolute Convergence Test Is $\sum |a_n|$ convergent? (Use step 2.)

   Alternating Test Is $a_n = (-1)^{n+1} b_n$, where (b_n) is non-negative, null and decreasing?

The following suggestions may also be helpful.

- If a_n is positive and includes $n!$ or c^n, then consider the Ratio Test.
- If a_n is positive and has dominant term n^p, then consider the Comparison Test or the Limit Comparison Test.
- If a_n includes a sine or cosine term, then use the fact that this term is bounded and consider the Comparison Test and the Absolute Convergence Test.
- If the terms of the series $\sum a_n$ are (eventually) only non-positive, apply step 2 of the strategy to the series $\sum(-a_n)$ and then use the Multiple Rule with $\lambda = -1$.
- If Strategy D13 doesn't give a result, then you could try using first principles by working directly with the sequence (s_n) of partial sums.

4 The exponential function

16. Alternative definition of the exponential function

Theorem D37

If $x \geq 0$, then

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

For $x < 0$,

$$e^x = (e^{-x})^{-1}.$$

17. Properties of e^x

Theorem D38

The number e is irrational.

Theorem D39

For any real numbers x and y, we have $e^{x+y} = e^x e^y$.

Unit D4 Continuity

1 Operations on functions

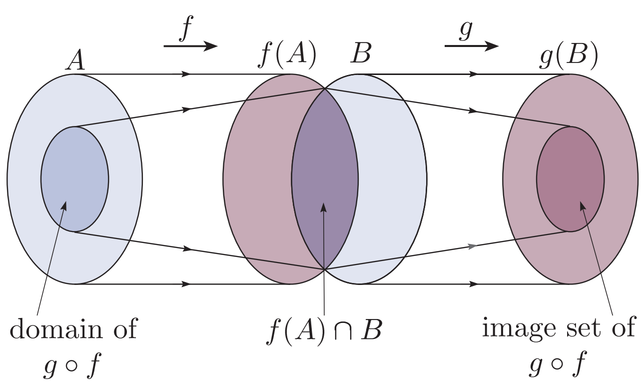
1. Convention for real functions

When a real function is specified *only by a rule*, it is understood that the domain of the function is the set of all real numbers for which the rule is applicable, and the codomain of the function is $\mathbb{R}$.

2. A function f is **defined on** a set I (usually an interval) if the domain of f contains the set I.

3. Let A and B be subsets of $\mathbb{R}$ and $f : A \longrightarrow \mathbb{R}$ and $g : B \longrightarrow \mathbb{R}$ be functions. Then

- the **sum** $f + g$ is the function with domain $A \cap B$ and rule
$$(f+g)(x) = f(x) + g(x)$$
- the **multiple** λf, for $\lambda \in \mathbb{R}$, is the function with domain A and rule
$$(\lambda f)(x) = \lambda f(x)$$
- the **product** fg is the function with domain $A \cap B$ and rule
$$(fg)(x) = f(x)g(x)$$
- the **quotient** f/g is the function with domain
$$A \cap B - \{x : g(x) = 0\}$$
and rule
$$(f/g)(x) = f(x)/g(x)$$
- the **composite** $g \circ f$ is the function with domain
$$\{x \in A : f(x) \in B\}$$
and rule
$$(g \circ f)(x) = g(f(x)).$$



4. A real function f defined on an interval I is

- **increasing** on I if
$$x_1 < x_2 \implies f(x_1) \le f(x_2), \quad \text{for } x_1, x_2 \in I$$
- **strictly increasing** on I if
$$x_1 < x_2 \implies f(x_1) < f(x_2), \quad \text{for } x_1, x_2 \in I$$
- **decreasing** on I if
$$x_1 < x_2 \implies f(x_1) \ge f(x_2), \quad \text{for } x_1, x_2 \in I$$
- **strictly decreasing** on I if
$$x_1 < x_2 \implies f(x_1) > f(x_2), \quad \text{for } x_1, x_2 \in I$$
- **monotonic** on I if f is either increasing on I or decreasing on I
- **strictly monotonic** on I if f is either strictly increasing on I or strictly decreasing on I.

A strictly monotonic function is one-to-one.

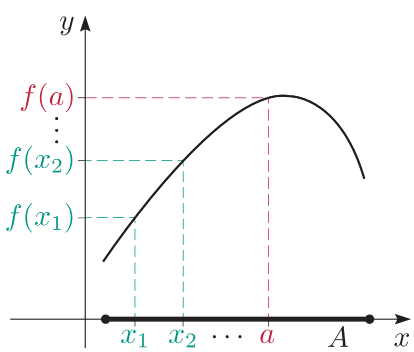
2 Continuous functions

5. Sequential definition of continuity

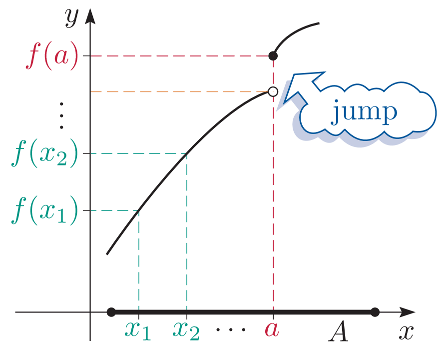
A function $f : A \longrightarrow \mathbb{R}$ is **continuous** at a point $a \in A$ if for each sequence (x_n) in A such that $x_n \to a$, we have $f(x_n) \to f(a)$.

We say that f is **continuous** (on A) if f is continuous at each point $a \in A$.

We say that f is **discontinuous** at a if f is not continuous at a point a in A.



f is continuous at a



f is discontinuous at a

Strategy D14 Continuous/discontinuous function

- To prove that a function $f : A \longrightarrow \mathbb{R}$ is *continuous* at the point $a \in A$:

  show that, if (x_n) is *any* sequence in A such that $x_n \to a$, then $f(x_n) \to f(a)$.
- To prove that a function $f : A \longrightarrow \mathbb{R}$ is *discontinuous* at the point $a \in A$:

  find *one* sequence (x_n) in A such that $x_n \to a$ but $f(x_n) \not\to f(a)$.

6. Rules for continuous functions

Theorem D40 Combination Rules

If f and g are continuous at a, then so are

Sum Rule	$f + g$	
Multiple Rule	λf,	for $\lambda \in \mathbb{R}$
Product Rule	fg	
Quotient Rule	f/g,	provided $g(a) \neq 0$.

Theorem D41

The following functions are continuous:

- any polynomial $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ (on its domain $\mathbb{R}$)
- any rational function $r(x) = p(x)/q(x)$, where p and q are polynomials (on its domain $\mathbb{R} - \{x : q(x) = 0\}$).

Theorem D42 Composition Rule

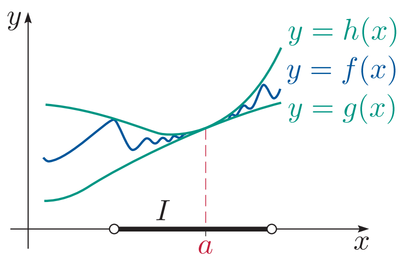
If f is continuous at a and g is continuous at $f(a)$, then $g \circ f$ is continuous at a.

Theorem D43 Squeeze Rule

Let f, g and h be defined on an open interval I and let $a \in I$. If

1. $g(x) \le f(x) \le h(x)$, for $x \in I$,
2. $g(a) = f(a) = h(a)$, and
3. g and h are continuous at a,

then f is also continuous at a.

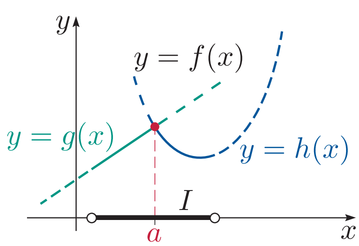


Theorem D44 Glue Rule

Let f be defined on an open interval I and let $a \in I$. If there are functions g and h such that

1. $f(x) = g(x)$, for $x \in I$, $x < a$,
   $f(x) = h(x)$, for $x \in I$, $x > a$,
2. $f(a) = g(a) = h(a)$, and
3. g and h are continuous at a,

then f is also continuous at a.



7. Continuity at a point is a local property because it depends only on the behaviour of the function near that point.

The restriction of a continuous function is continuous.

8. Trigonometric functions

Theorem D45 Sine Inequality

$$\sin x \le x, \quad \text{for } 0 \le x \le \frac{\pi}{2}.$$

Corollary D46

$$|\sin x| \le |x|, \quad \text{for } x \in \mathbb{R}.$$

Theorem D47

The trigonometric functions sine, cosine and tangent are continuous.

9. The exponential function

Theorem D48 Exponential Inequalities

(a) $e^x \ge 1 + x, \quad \text{for } x \ge 0$

(b) $e^x \le \dfrac{1}{1-x}, \quad \text{for } 0 \le x < 1.$

Corollary D49

$$1 + x \le e^x \le \frac{1}{1-x}, \quad \text{for } |x| < 1.$$

Theorem D50

The exponential function is continuous.

10. The following theorem collects together Theorems D41, D47 and D50 along with two additional continuous functions.

Theorem D51 Basic continuous functions

The following functions are continuous:

- polynomials and rational functions
- $f(x) = |x|$
- $f(x) = \sqrt{x}$
- the trigonometric functions sine, cosine and tangent
- the exponential function.

3 Properties of continuous functions

11. Intermediate Value Theorem

Theorem D52 Intermediate Value Theorem

Let f be a function continuous on $[a, b]$ such that $f(a) \neq f(b)$, and let k be any number lying between $f(a)$ and $f(b)$. Then there exists a number c in (a, b) such that

$$f(c) = k.$$

Theorem D53 (special case)

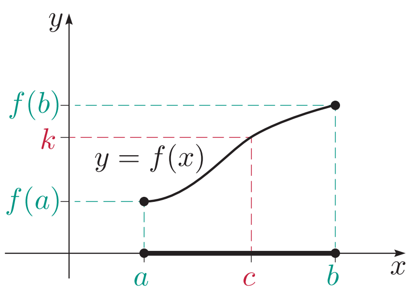
Let f be a function continuous on $[a, b]$ and suppose that

$$f(a) < 0 < f(b).$$

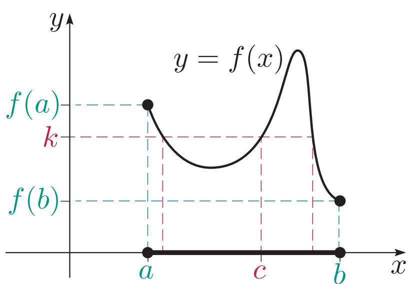
Then there exists a number c in (a, b) such that

$$f(c) = 0.$$

There may be more than one possible value of c such that $f(c) = k$, or $f(c) = 0$, respectively.



$f(a) < f(b)$

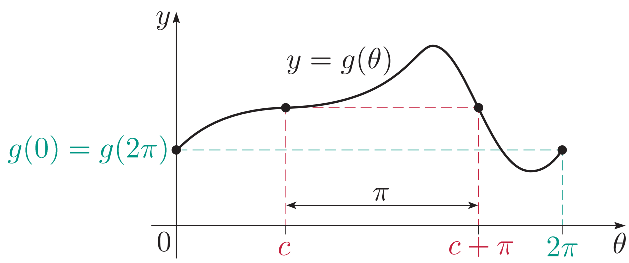


$f(a) > f(b)$

Theorem D57 Antipodal Points Theorem

If $g : [0, 2\pi] \longrightarrow \mathbb{R}$ is a continuous function and $g(0) = g(2\pi)$, then there exists a number c in $[0, \pi]$ such that

$$g(c) = g(c + \pi).$$



12. Bisection method

Given that there is a solution to $f(c) = k$ in an interval (a, b), the **bisection method** involves repeatedly bisecting the interval and determining the values of f at the bisection points, in order to find shorter and shorter intervals in which the solution must lie.

13. A **zero** of a function f is a real number c such that $f(c) = 0$. We also say that f **vanishes** at c.

14. Locating zeros of polynomials

A real polynomial of degree n has at most n zeros (distinct roots).

Theorem D54

Let

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

for $x \in \mathbb{R}$, where $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$.

Then all the zeros of p (if there are any) lie in the open interval $(-M, M)$, where

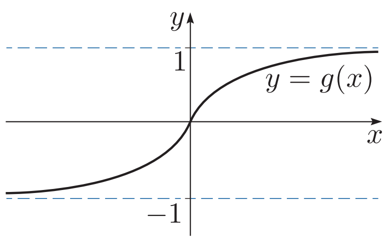
$$M = 1 + \max\{|a_{n-1}|, \ldots, |a_1|, |a_0|\}.$$

15. Let f be a function with domain A. Then

- f has **maximum value** $f(c)$ in A if $c \in A$ and
$$f(x) \le f(c), \quad \text{for } x \in A$$
- f has **minimum value** $f(c)$ in A if $c \in A$ and
$$f(c) \le f(x), \quad \text{for } x \in A$$
- f is **bounded** on A if, for some $M \in \mathbb{R}$,
$$|f(x)| \le M, \quad \text{for } x \in A.$$

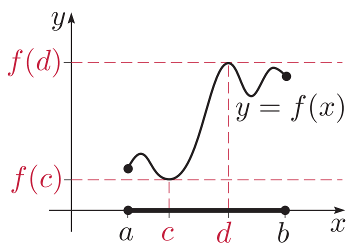
An **extreme value** is either a maximum or a minimum value.

A function can be bounded without having a maximum value or a minimum value.



Theorem D55 Extreme Value Theorem

Let f be a function continuous on $[a,b]$. Then there exist numbers c and d in $[a,b]$ such that
$$f(c) \le f(x) \le f(d), \quad \text{for } x \in [a,b].$$



Corollary D56 Boundedness Theorem

Let f be a function continuous on $[a,b]$. Then there exists a number M such that
$$|f(x)| \le M, \quad \text{for } x \in [a,b].$$

4 Inverse functions

16. Finding inverse functions

Theorem D58 Inverse Function Rule

Let $f: I \longrightarrow J$, where I is an interval and J is the image set $f(I)$, be a function such that

1. f is strictly increasing on I
2. f is continuous on I.

Then J is an interval and f has an inverse function $f^{-1}: J \longrightarrow I$ such that

1′. f^{-1} is strictly increasing on J

2′. f^{-1} is continuous on J.

The interval I may be of *any* type: open or closed, half-open, bounded or unbounded.

There is another version of the Inverse Function Rule with 'strictly increasing' replaced by 'strictly decreasing'.

The graph of $f^{-1}(x)$ is obtained by reflecting the graph of $f(x)$ in the line $y = x$.

Strategy D15 Inverse function

To prove that a function $f: I \longrightarrow J$, where I is an interval with endpoints a and b, has a continuous inverse function $f^{-1}: J \longrightarrow I$, do the following:

1. Show that f is strictly increasing on I.
2. Show that f is continuous on I.
3. Determine the endpoint c of J corresponding to the endpoint a of I as follows:
   - if $a \in I$, then $c = f(a)$ and $c \in J$
   - if $a \notin I$, then $f(a_n) \to c$ and $c \notin J$, where (a_n) is a monotonic sequence in I such that $a_n \to a$.

   Determine the endpoint d of J, corresponding to the endpoint b of I in a similar way.

There is a corresponding version of this strategy for a function that is strictly decreasing.

17. Inverse functions

Sketches of inverse functions are included in the Quick reference section on page 130.

- ***n*th root function** For any positive integer $n \geq 2$, the function
  $$f(x) = x^n \quad (x \in [0, \infty)),$$
  has a strictly increasing continuous inverse function
  $$f^{-1}(x) = \sqrt[n]{x},$$
  with domain $[0, \infty)$ and image set $[0, \infty)$, called the ***n*th root function**.
- **log** The function
  $$f(x) = e^x \quad (x \in \mathbb{R})$$
  has a strictly increasing continuous inverse function f^{-1}, with domain $(0, \infty)$ and image set $\mathbb{R}$, called **log** (or **ln**).
  For all $x, y > 0$,
  $$\log(xy) = \log x + \log y.$$

Inverse trigonometric functions

- $\mathbf{sin^{-1}}$ The function
  $$f(x) = \sin x \quad \left(x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]\right)$$
  has a strictly increasing continuous inverse function, with domain $[-1, 1]$ and image set $[-\pi/2, \pi/2]$, called $\mathbf{sin^{-1}}$.
- $\mathbf{cos^{-1}}$ The function
  $$f(x) = \cos x \quad (x \in [0, \pi])$$
  has a strictly decreasing continuous inverse function, with domain $[-1, 1]$ and image set $[0, \pi]$, called $\mathbf{cos^{-1}}$.
- $\mathbf{tan^{-1}}$ The function
  $$f(x) = \tan x \quad \left(x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)\right)$$
  has a strictly increasing continuous inverse function, with domain $\mathbb{R}$ and image set $(-\pi/2, \pi/2)$, called $\mathbf{tan^{-1}}$.

Inverse hyperbolic functions

- $\mathbf{sinh^{-1}}$ The function
  $$f(x) = \sinh x = \tfrac{1}{2}(e^x - e^{-x}) \quad (x \in \mathbb{R})$$
  has a strictly increasing continuous inverse function f^{-1}, with domain $\mathbb{R}$ and image set $\mathbb{R}$, called $\mathbf{sinh^{-1}}$.
- $\mathbf{cosh^{-1}}$ The function
  $$f(x) = \cosh x = \tfrac{1}{2}(e^x + e^{-x}) \quad (x \in [0, \infty))$$
  has a strictly increasing continuous inverse function f^{-1}, with domain $[1, \infty)$ and image set $[0, \infty)$, called $\mathbf{cosh^{-1}}$.
- $\mathbf{tanh^{-1}}$ The function
  $$f(x) = \tanh x = \frac{\sinh x}{\cosh x} \quad (x \in \mathbb{R})$$
  has a strictly increasing continuous inverse function f^{-1}, with domain $(-1, 1)$ and image set $\mathbb{R}$, called $\mathbf{tanh^{-1}}$.

18. Defining exponential functions

If $a > 0$, then

$$a^x = e^{x \log a} \quad (x \in \mathbb{R}).$$

Theorem D59

(a) If $a > 0$, then the function
$$x \longmapsto a^x = e^{x \log a} \quad (x \in \mathbb{R})$$
is continuous.

(b) If $a, b > 0$ and $x, y \in \mathbb{R}$, then
- $a^x b^x = (ab)^x$
- $a^x a^y = a^{x+y}$
- $(a^x)^y = a^{xy}$.

Book E Group theory 2

Unit E1 Cosets and normal subgroups

2 Matrix groups

Throughout, all matrix entries are real numbers.

1. Standard matrix groups

> **Theorem E1**
>
> The set of invertible 2×2 matrices is a group under matrix multiplication.

The **general linear group of degree** 2, denoted by GL(2), is the group in Theorem E1:

$$\mathrm{GL}(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R},\ ad - bc \neq 0 \right\}.$$

The following are subgroups of GL(2).

- The **special linear group of degree** 2, denoted by SL(2), whose elements are all 2×2 matrices with determinant 1:
$$\mathrm{SL}(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R},\ ad - bc = 1 \right\}.$$
- The group L of invertible 2×2 lower triangular matrices:
$$L = \left\{ \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} : a, c, d \in \mathbb{R},\ ad \neq 0 \right\}.$$
- The group U of invertible 2×2 upper triangular matrices:
$$U = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, b, d \in \mathbb{R},\ ad \neq 0 \right\}.$$
- The group D of invertible 2×2 diagonal matrices:
$$D = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} : a, d \in \mathbb{R},\ ad \neq 0 \right\}.$$

The **general linear group of degree** n, denoted by GL(n), is the group of all invertible $n \times n$ matrices under matrix multiplication.

The **special linear group of degree** n, denoted by SL(n), is the group of all $n \times n$ matrices with determinant 1 under matrix multiplication.

4 Cosets

2. Left cosets and right cosets

Let H be a subgroup of a group G, and let g be an element of G.

The **left coset** gH of H is given by

$$gH = \{gh : h \in H\}.$$

The **right coset** Hg of H is given by

$$Hg = \{hg : h \in H\}.$$

These are the subsets of G obtained by composing each element of H with g on the left, or right, respectively.

> **Theorems E2 and E4**
>
> Let H be a subgroup of a group G.
>
> - The distinct left cosets of H in G form a partition of G.
> - The distinct right cosets of H in G form a partition of G.

The partition of a group G into left cosets of a subgroup H may or may not be the same as the partition of G into right cosets of H.

> **Propositions E3 and E5 Properties of left and right cosets**
>
> Let H be a subgroup of a group G.
>
> (a) The element g lies in the left coset gH, for each $g \in G$.
>
> (b) One of the left cosets of H in G is H itself.
>
> (c) Any two left cosets g_1H and g_2H are either the same set or are disjoint.
>
> (d) If H is finite, then each left coset gH has the same number of elements as H.
>
> Properties (a)–(d) hold if 'left' is replaced by 'right' throughout, and all left cosets 'gH' are replaced by right cosets 'Hg'.

3. Cosets in infinite groups

A subgroup of an infinite group G may have a finite number of left cosets in G, or infinitely many. The same holds for right cosets.

4. Relationship between left and right cosets

Theorem E6

Let H be a subgroup of a group G.

(a) If every element in the partition of G into left cosets of H is replaced by its inverse, then the result is the partition of G into right cosets of H.

(b) The same is true if the words 'left' and 'right' are interchanged.

Corollary E7

Let H be a subgroup of a group G. Then the number of distinct left cosets of H in G is equal to the number of distinct right cosets of H in G (or there may be infinitely many of each).

***Warning*:** Theorem E6 does not say that if H is a subgroup of a group G and g is an element of G then replacing every element of the left coset gH by its inverse gives the right coset Hg. This procedure gives a right coset of H in G, but it may not be the right coset Hg.

5. Index of a subgroup in a group

The **index** of a subgroup H in a group G is the number of distinct left (or right) cosets of H in G.

We say that H has **infinite index** in G if H has infinitely many left (or right) cosets in G.

Proposition E8

Let H be a subgroup of a finite group G. Then the index of H in G is $|G|/|H|$.

6. Cosets in abelian groups

If G is an abelian group, H is a subgroup of G and g is an element of G, then the left coset of H containing g is the same as the right coset of H containing g, and we refer to it as the coset of H containing g.

7. Cosets in additive groups

Let H be a subgroup of an additive group $(G, +)$, and let g be an element of G. We denote the coset of H containing g by $g + H$; thus

$$g + H = \{g + h : h \in H\}.$$

(Since G is abelian, the left cosets and the right cosets of H in G are the same.)

8. Partitioning a finite group into left cosets or right cosets

Strategies E1 and E2 Finding cosets

- To partition a finite group G into left cosets of a subgroup H, do the following.
  1. Take H as the first left coset.
  2. Choose any element $g \in G$ not yet assigned to a left coset and determine the left coset gH to which g belongs.
  3. Repeat step 2 until every element of G has been assigned to a left coset.
- To partition a finite group G into right cosets of a subgroup H, follow the steps above with 'left' replaced by 'right' throughout, and the left coset gH replaced by the right coset Hg.

These strategies can also be used if G is infinite and H has only finitely many left cosets and right cosets in G.

***Alternative*:** When the inverses of the elements of G and the left cosets of H in G are known, it is easier to use Theorem E6 than Strategy E2 to find the right cosets of H in G. Similarly, Theorem E6 can be used to find the left cosets when the inverses and right cosets are known.

5 Normal subgroups

9. A subgroup H of a group G is a **normal subgroup** of G if the partition of G into left cosets of H is the same as the partition of G into right cosets of H.

We also say that H is **normal in** G.

Theorem E9

The following are normal subgroups of any group G.

(a) The trivial subgroup $\{e\}$.

(b) The whole group G.

Theorem E10

In an abelian group, every subgroup is normal.

Theorem E11

Every subgroup of index 2 in a group is a normal subgroup of the group.

Corollary E12

For each positive integer n, the alternating group A_n is a normal subgroup of the symmetric group S_n.

Proposition E13

Let H be a subgroup of a group G. Then H is normal in G if and only if

$$gH = Hg$$

for each element $g \in G$.

***Warning*:** The equation $gH = Hg$ in Proposition E13 means that the sets gH and Hg contain the same elements; it does not mean that $gh = hg$ for all $h \in H$.

If N is a normal subgroup of a group G, then since the left cosets of N are the same as the right cosets of N we refer to the cosets of N in G.

Unit E2 Quotient groups and conjugacy

1 Quotient groups

1. **Set composition** in a group G is the binary operation $\cdot$ defined on the set of subsets of G by

$$X \cdot Y = \{xy : x \in X,\ y \in Y\}$$

for all subsets X and Y of G.

Thus, if X and Y are subsets of G then $X \cdot Y$ is the subset of G obtained by composing each element of X with each element of Y, in that order.

Theorem E14

Let N be a normal subgroup of a group G. Then, for all $x, y \in G$,

$$xN \cdot yN = (xy)N.$$

Theorem E14 tells us that, for a normal subgroup N, the set that we obtain by composing two cosets of N using set composition is always a coset of N.

Theorem E15

Let N be a normal subgroup of a group G. Then the set of cosets of N in G, with the binary operation of set composition, is a group.

The group obtained in Theorem E15 is the **quotient group** of G by N (see **2.** on page 96).

2. Quotient groups

Let N be a normal subgroup of a group G. The **quotient group** or **factor group** G/N of G by N is the group of cosets of N in G under set composition.

- Set composition of elements of G/N is given by
  $$xN \cdot yN = (xy)N \quad \text{for each } x, y \in G.$$
  If G is additive, then this is written as
  $$(x+N)+(y+N) = (x+y)+N \quad \text{for each } x, y \in G.$$
- The identity of G/N is N.
- For each $x \in G$, the inverse of xN is $x^{-1}N$. If G is additive, the inverse of $x+N$ is written as $-x+N$.
- If G is finite, then $|G/N| = |G|/|N|$.

Strategy E3 Identifying G/N

To find a standard group isomorphic to a finite quotient group G/N where N is a normal subgroup of the group G, do the following.

1. Determine the cosets of N in G, by repeatedly choosing an element x of G not yet assigned to a coset and finding the coset xN (or $x+N$, if G is additive) until all the elements of G have been assigned to cosets.
2. Construct the group table of G/N by composing each pair of cosets using the rule
   $$xN \cdot yN = (xy)N$$
   (or
   $$(x+N)+(y+N) = (x+y)+N$$
   if G is additive).
   Make sure to use just one way to write each coset.
3. By inspection of the group table, and possibly using the tables of isomorphism classes in the Quick reference section on page 137, identify a standard group isomorphic to G/N.

Let N be a normal subgroup of a group G.

- If G is abelian, then G/N is abelian.
- If G is non-abelian, then G/N may be either abelian or non-abelian.

3. Quotient groups of $(\mathbb{Z}, +)$

Since all subgroups of an abelian group are normal, each subgroup $n\mathbb{Z}$ of $\mathbb{Z}$ is normal.

Theorem E16

For each integer $n \geq 2$, the quotient group $\mathbb{Z}/n\mathbb{Z}$ is isomorphic to the group $\mathbb{Z}_n$, and the following mapping is an isomorphism:
$$\begin{aligned} \phi : \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}_n \\ a + n\mathbb{Z} &\longmapsto a, \quad \text{for } a = 0, 1, 2, \ldots, n-1. \end{aligned}$$

Thus each quotient group $\mathbb{Z}/n\mathbb{Z}$ has finite order n.

4. The **fractional part**, $\operatorname{frac}(x)$, of a real number x is

$$\operatorname{frac}(x) = x - \lfloor x \rfloor,$$

where $\lfloor x \rfloor$ is the integer part of x (the largest integer that is less than or equal to x). Thus $\operatorname{frac}(x)$ is always a number in the interval $[0,1)$.

The binary operation $+_1$ is defined on $[0,1)$ by

$$x +_1 y = \operatorname{frac}(x+y).$$

The interval $[0,1)$ is a group under $+_1$.

5. The quotient group $\mathbb{R}/\mathbb{Z}$

We can express any coset $y + \mathbb{Z}$ of $\mathbb{Z}$ in $\mathbb{R}$ in the form $x + \mathbb{Z}$ where $x \in [0,1)$ by taking $x = \operatorname{frac}(y)$.

Theorem E17

The quotient group $\mathbb{R}/\mathbb{Z}$ is isomorphic to the group $([0,1), +_1)$, and the following mapping is an isomorphism:
$$\begin{aligned} \phi : \mathbb{R}/\mathbb{Z} &\longrightarrow [0,1) \\ x + \mathbb{Z} &\longmapsto x, \quad \text{for } x \in [0,1). \end{aligned}$$

Thus the group $\mathbb{R}/\mathbb{Z}$ is infinite. It is given by

$$\mathbb{R}/\mathbb{Z} = \{x + \mathbb{Z} : x \in [0,1)\},$$

with binary operation

$$(x+\mathbb{Z}) + (y+\mathbb{Z}) = (x +_1 y) + \mathbb{Z}.$$

2 Conjugacy

6. An element y in a group G is a **conjugate** in G of an element x in G if there exists an element g in G such that

$$y = gxg^{-1}.$$

We say that:

- g **conjugates** x to y
- y is the **conjugate** of x by g
- g is a **conjugating element**
- x and y are **conjugates**, or **conjugate elements**, or they are **conjugate** in G.

Note that g^{-1} conjugates y to x.

Lemma E20

Let x, y and g be elements of a group, and suppose that $y = gxg^{-1}$. Then $y^n = gx^n g^{-1}$ for all positive integers n.

Theorem E21

Let x and y be conjugate elements in a group G. Then either x and y have the same finite order, or they both have infinite order.

***Warning*:** The converse of Theorem E21 is not true: group elements of the same order are not necessarily conjugate.

7. Conjugacy classes

Let G be a group, and let $x \in G$. The **conjugacy class** of x in G is the set of all elements of G that are conjugate to x. That is, it is the set

$$\{gxg^{-1} : g \in G\}.$$

Theorem E22

Let G be a group. Then the relation 'is a conjugate of' is an equivalence relation on the set of elements of G.

Corollary E23

In any group, the distinct conjugacy classes form a partition of the group.

The **conjugacy classes of a group** G are the distinct conjugacy classes of the elements of G.

Proposition E24

Let G be a group with identity element e. Then $\{e\}$ is a conjugacy class of G.

Theorem E25

In an abelian group, each conjugacy class contains a single element.

Proposition E26

Let H be a subgroup of a group G, and let x and y be elements of H.

(a) If x and y are conjugate in H, then they are also conjugate in G.

(b) If x and y are conjugate in G, then they may or may not be conjugate in H.

Theorem E27

In any finite group G, the number of elements in each conjugacy class divides the order of G.

8. The group $S(\triangle)$ has three conjugacy classes:

$$\{e\}, \quad \{a, b\}, \quad \{r, s, t\}.$$

The group $S(\square)$ has five conjugacy classes:

$$\{e\}, \quad \{b\}, \quad \{a, c\}, \quad \{r, t\}, \quad \{s, u\}.$$

9. Conjugacy of permutations is covered in **25.** to **30.** on pages 47 and 48.

3 Normal subgroups and conjugacy

10. Checking the **normality** of a subgroup means checking whether it is normal.

Theorem E28

Let G be a group and let H be a subgroup of G. Then H is a normal subgroup of G if and only if

$$ghg^{-1} \in H$$

for each $h \in H$ and each $g \in G$.

Theorem E28 can also be expressed as follows.

Let G be a group and let H be a subgroup of G. Then H is a normal subgroup of G if and only if

$$gHg^{-1} \subseteq H \quad \text{for each } g \in G.$$

11. Conjugate subgroups

Let H be a subgroup of a group G, and let g be any element of G. Then

$$gHg^{-1} \text{ denotes the set } \{ghg^{-1} : h \in H\}.$$

That is, gHg^{-1} is the set obtained by conjugating every element of H by the element g.

Theorem E29

Let H be a subgroup of a group G and let g be any element of G. Then the subset gHg^{-1} is a subgroup of G.

We say that:

- gHg^{-1} is the **conjugate subgroup** of H by g
- gHg^{-1} is a **conjugate subgroup** of H in G
- g **conjugates** H to gHg^{-1}
- H and gHg^{-1} are **conjugate subgroups** in G.

Note that g^{-1} conjugates gHg^{-1} to H.

Proposition E30

Let H and K be conjugate subgroups in a group G. Then either H and K have the same finite order, or they both have infinite order.

12. Conjugate subgroups and normality

Theorem E31

Let G be a group and let H be a subgroup of G. Then H is a normal subgroup of G if and only if

$$gHg^{-1} = H \quad \text{for each } g \in G.$$

13. Normal subgroups and conjugacy classes

Theorem E32

Let G be a group and let H be a subgroup of G. Then H is a normal subgroup of G if and only if

H is a union of conjugacy classes of G.

***Warning*:** The condition in Theorem E32 that H is a subgroup of G is essential.

Strategy E5 Finding normal subgroups using conjugacy clases

To find all the normal subgroups of a finite group G, do the following.

1. Partition G into conjugacy classes.
2. Find all the unions of conjugacy classes that include the class $\{e\}$ and whose total number of elements is a divisor of $|G|$.
3. Determine whether each such union of conjugacy classes is a subgroup of G: any union that is a subgroup is a normal subgroup of G.

14. The normal subgroups of $S(\triangle)$ are

$$\{e\}, \quad \{e, a, b\}, \quad S(\triangle).$$

The normal subgroups of $S(\square)$ are

$$\{e\}, \quad \{e, b\}, \quad \{e, a, b, c\},$$
$$\{e, b, r, t\}, \quad \{e, b, s, u\}, \quad S(\square).$$

The normal subgroups of S_4 are

$$\{e\}, \quad \{e, (1\ 2)(3\ 4), (1\ 3)(2\ 4), (1\ 4)(2\ 3)\},$$
$$A_4, \quad S_4.$$

15. Characterising normality

A property **characterises** normal subgroups if every normal subgroup has this property and every subgroup with this property is a normal subgroup.

Theorem E33 below collects together Proposition E13 and Theorems E28, E31 and E32.

> **Theorem E33**
>
> A subgroup H of a group G is normal in G if and only if it has any one of the following equivalent properties.
>
> **Property A** $gH = Hg$ for each $g \in G$.
>
> **Property B** $ghg^{-1} \in H$ for each $h \in H$ and each $g \in G$.
>
> **Property C** $gHg^{-1} = H$ for each $g \in G$.
>
> **Property D** H is a union of conjugacy classes of G.

Any of the conditions in Theorem E33 can be used to prove that a subgroup is normal, or show that it is not normal.

Property A is useful when we know the partition into left cosets and the partition into right cosets.

Property B is useful in many general situations.

- To show that a subgroup H of a group G is normal in G, take a general element $g \in G$ and a general element $h \in H$, and show that the conjugate ghg^{-1} belongs to H.
- To show that a subgroup H of a group G is not normal in G, find one element $h \in H$ and one element $g \in G$ such that the conjugate ghg^{-1} does not belong to H.

Property C may be helpful when we have knowledge about conjugate subgroups.

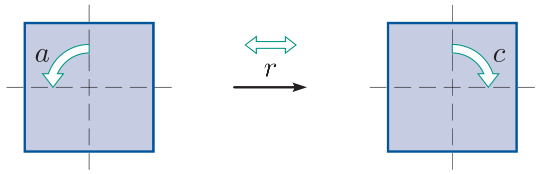
Property D is particularly useful when we know the conjugacy classes. However, remember that not every union of conjugacy classes of G is a normal subgroup of G, only those that are also subgroups of G.

4 Conjugacy in symmetry groups

16. Conjugate symmetries

- Two symmetries x and y of a figure F are conjugate in $S(F)$ if and only if there is a symmetry g of F that transforms a diagram illustrating x into a diagram illustrating y (when we ignore any labels).
- If such a symmetry g exists, then $y = g \circ x \circ g^{-1}$.

For example, the symmetries a and c of the square are conjugate in $S(\square)$ because the symmetry r of the square transforms a diagram illustrating a into a diagram illustrating c.



Thus $c = r \circ a \circ r^{-1}$.

By the first bullet point above, conjugate symmetries are those of the same geometric type.

17. The **fixed point set**, $\operatorname{Fix} g$, of a symmetry g of a figure F is

$$\operatorname{Fix} g = \{P \in F : g(P) = P\}.$$

> **Theorem E34**
>
> Let x and g be symmetries of a figure F, and let the fixed point set of x be L. Then the fixed point set of $g \circ x \circ g^{-1}$ is $g(L)$.

Thus if x and y are symmetries of a figure F, then any symmetry g of F that does not map $\operatorname{Fix} x$ to $\operatorname{Fix} y$ does not conjugate x to y.

***Warning*:** If a symmetry g maps $\operatorname{Fix} x$ to $\operatorname{Fix} y$ then there is no guarantee that it conjugates x to y: it may or may not do this.

If there is *no* symmetry in $S(F)$ that maps $\operatorname{Fix} x$ to $\operatorname{Fix} y$, then x and y are not conjugate.

> **Theorem E35**
>
> A direct symmetry cannot be conjugate to an indirect symmetry in a symmetry group.

18. If we represent a symmetry group $S(F)$ as a subgroup of a symmetric group S_n (by labelling the vertices of F, for example), then any symmetries that are conjugate in $S(F)$ must also be conjugate in S_n, and hence must have the same cycle structure.

Strategy E6 Finding conjugacy classes

To determine the conjugacy classes of a finite symmetry group $S(F)$, do the following.

1. Represent $S(F)$ as a group of permutations.
2. Partition $S(F)$ by cycle structure.
3. For each cycle structure class, determine whether all the symmetries in the class are conjugate to each other, or whether the class splits into two or more conjugacy classes.

   The following can help you do this.

   - Two symmetries x and y are conjugate in $S(F)$ if and only if there is a symmetry g of F that transforms a diagram illustrating x into a diagram illustrating y.
   - If x and y are conjugate in a subgroup H of $S(F)$, then they are also conjugate in $S(F)$.
   - If x and y are not conjugate in a group G that has $S(F)$ as a subgroup, then they are not conjugate in $S(F)$.
   - If the fixed point set of x is L, then the fixed point set of $g \circ x \circ g^{-1}$ is $g(L)$.
   - A direct symmetry and an indirect symmetry are not conjugate.
   - Renaming method: To find the conjugate $g \circ x \circ g^{-1}$, replace each symbol in the cycle form of x by its image under g.
   - The number of elements in each conjugacy class divides $|S(F)|$.

Unit E3 Homomorphisms

Throughout, e_G and e_H denote the identity elements of groups $(G, \circ)$ and $(H, *)$, respectively.

1 Isomorphisms and homomorphisms

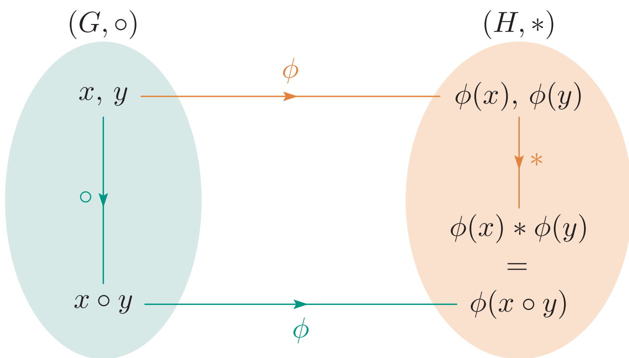
1. Homomorphisms

A **homomorphism** is a mapping $\phi : (G, \circ) \longrightarrow (H, *)$, where $(G, \circ)$ and $(H, *)$ are groups, that has the property

$$\phi(x \circ y) = \phi(x) * \phi(y) \quad \text{for all } x, y \in G.$$

This property is the **homomorphism property**.

The homomorphism property means that composites are preserved under a homomorphism.



2. An **isomorphism** is a homomorphism that is one-to-one and onto.

Proposition E36

Let $(G, \circ)$ and $(H, *)$ be groups. If ϕ is an isomorphism from $(G, \circ)$ to $(H, *)$, then ϕ^{-1} is an isomorphism from $(H, *)$ to $(G, \circ)$.

An **automorphism** of a group is an isomorphism from the group to itself.

3. Let k be any integer and let n be any integer with $n \geq 2$.

The **least residue of k modulo n** is the integer in $\mathbb{Z}_n$ that is congruent to k modulo n; it is the remainder of k on division by n. We denote it by $k_{(\text{mod } n)}$.

Proposition E37

For any integer $n \geq 2$, the following mapping is a homomorphism:

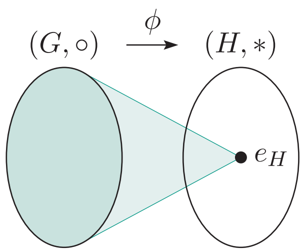
$$\begin{aligned} \phi : (\mathbb{Z}, +) &\longrightarrow (\mathbb{Z}_n, +_n) \\ k &\longmapsto k_{(\text{mod } n)}. \end{aligned}$$

4. Trivial homomorphism

Proposition E38

Let $(G, \circ)$ and $(H, *)$ be groups. Then the following mapping is a homomorphism:

$$\begin{aligned} \phi : (G, \circ) &\longrightarrow (H, *) \\ x &\longmapsto e_H. \end{aligned}$$



The homomorphism in Proposition E38 is the **trivial homomorphism**; it maps every element of $(G, \circ)$ to the identity element of $(H, *)$.

5. Linear transformations as homomorphisms

Proposition E39

Let V and W be vector spaces and let $t : V \longrightarrow W$ be a linear transformation. Then t is a homomorphism from the group $(V, +)$ to the group $(W, +)$.

Thus some mappings can immediately be recognised as homomorphisms. For example, any mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$ of the form

$$(x, y) \longmapsto (ax + by, cx + dy),$$

where $a, b, c, d \in \mathbb{R}$, is a homomorphism from $(\mathbb{R}^2, +)$ to $(\mathbb{R}^2, +)$, because it has a matrix representation and so is a linear transformation. (See Theorem C41 in **10.** on page 66.)

6. Properties of homomorphisms

An isomorphism preserves *all* of the structure of its domain group.

A homomorphism preserves *some* of the structure of its domain group.

As detailed below, all homomorphisms preserve:

- composites of any finite number of elements
- the identity
- inverses
- powers
- conjugates.

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism.

Proposition E40

If $x_1, x_2, \ldots, x_n$ are any elements of G, then

$$\phi(x_1 \circ x_2 \circ \cdots \circ x_n) \\ = \phi(x_1) * \phi(x_2) * \cdots * \phi(x_n).$$

Proposition E41

$$\phi(e_G) = e_H.$$

Proposition E42

For all $x \in G$,

$$\phi(x^{-1}) = (\phi(x))^{-1}.$$

Proposition E43

For all $x \in G$ and all $n \in \mathbb{Z}$,

$$\phi(x^n) = (\phi(x))^n.$$

Proposition E46

For all $x, y \in G$, if x and y are conjugate in $(G, \circ)$, then $\phi(x)$ and $\phi(y)$ are conjugate in $(H, *)$.

Homomorphisms do not in general preserve the orders of elements.

Theorem E44

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism and let x be an element of finite order in G. Then the order of $\phi(x)$ is finite and divides the order of x.

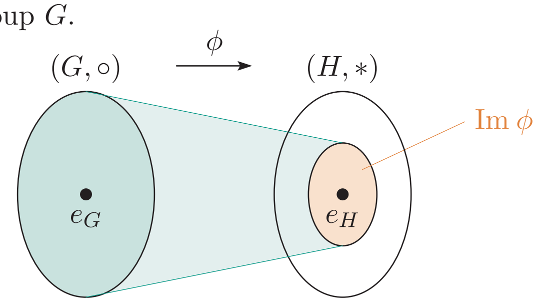
2 Images and kernels

7. Image of a homomorphism

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism. The **image** (or **image set**) of ϕ is

$$\operatorname{Im} \phi = \{\phi(g) : g \in G\}.$$

It is the set of elements of the codomain group H that are images of elements in the domain group G.



A homomorphism $\phi : (G, \circ) \longrightarrow (H, *)$ is onto if and only if $\operatorname{Im} \phi = H$.

Theorem E47

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism. Then $\operatorname{Im} \phi$ is a subgroup of $(H, *)$.

8. Further properties of homomorphisms

Proposition E48

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a one-to-one homomorphism, and let θ be the mapping obtained from ϕ by shrinking the codomain of ϕ to its subgroup $\operatorname{Im} \phi$. Then θ is an isomorphism, and hence $(G, \circ) \cong \operatorname{Im} \phi$.

Thus every one-to-one homomorphism ϕ preserves *all* of the structure of the domain group, in the subgroup $\operatorname{Im} \phi$ of the codomain group.

Theorem E49

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism.

(a) If $(G, \circ)$ is abelian, then $(\operatorname{Im} \phi, *)$ is abelian.

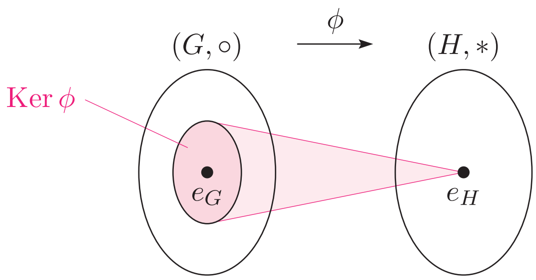
(b) If $(G, \circ)$ is cyclic, then $(\operatorname{Im} \phi, *)$ is cyclic. In particular, if $(G, \circ)$ is generated by a, then $(\operatorname{Im} \phi, *)$ is generated by $\phi(a)$.

9. Kernel of a homomorphism

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism. The **kernel** of ϕ is

$$\operatorname{Ker} \phi = \{g \in G : \phi(g) = e_H\}.$$

It is the set of elements of the domain group G that are mapped by ϕ to e_H.



The definition of a kernel, unlike the definition of an image, applies only to homomorphisms, not to functions in general. This is because the codomain of a function need not contain an identity element.

Theorems E50 and E51

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism. Then $\operatorname{Ker} \phi$ is a normal subgroup of $(G, \circ)$.

Theorem E52

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism. Then ϕ is one-to-one if and only if $\operatorname{Ker} \phi = \{e_G\}$.

10. Kernels of homomorphisms and normal subgroups

Theorem E53

Let K be a subgroup of a group $(G, \circ)$. Then K is normal in G if and only if K is the kernel of a homomorphism with domain group G.

Thus kernels of homomorphisms and normal subgroups are essentially the same objects.

3 The First Isomorphism Theorem

11. Cosets of the kernel of a homomorphism

The kernel of a homomorphism ϕ is a normal subgroup of the domain group $(G, \circ)$, so its left cosets are the same as its right cosets and the quotient group $G/\operatorname{Ker}\phi$ can be formed.

> **Theorem E54**
>
> Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism, and let x and y be any elements of G. Then
>
> x and y have the same image under ϕ
>
> if and only if
>
> x and y lie in the same coset of $\operatorname{Ker}\phi$ in G.

In other words, the cosets of $\operatorname{Ker}\phi$ in $(G, \circ)$ are the sets of elements of G with the same image under ϕ.

12. The First Isomorphism Theorem

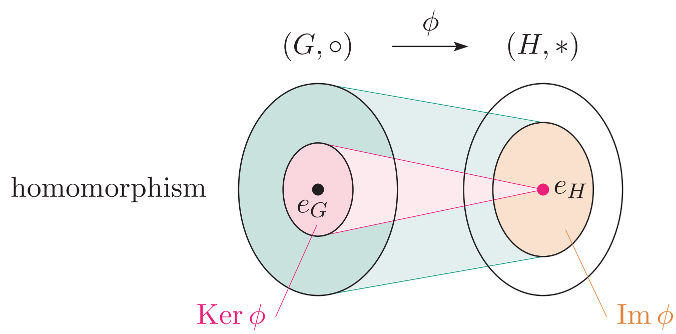
The First Isomorphism Theorem says that for *any* homomorphism ϕ, with domain group G, the quotient group $G/\operatorname{Ker}\phi$ is isomorphic to the image group $\operatorname{Im}\phi$.

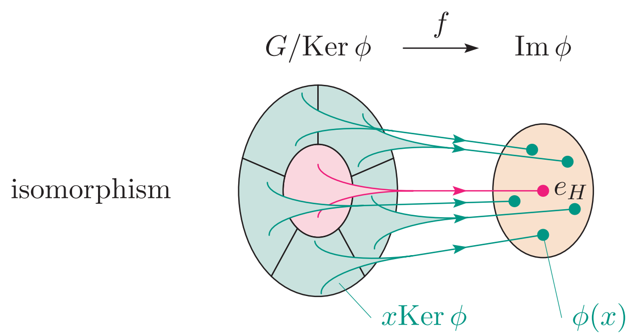
> **Theorem E55 First Isomorphism Theorem**
>
> Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism. Then the mapping
>
> $$\begin{aligned} f : G/\operatorname{Ker}\phi &\longrightarrow \operatorname{Im}\phi \\ x\operatorname{Ker}\phi &\longmapsto \phi(x) \end{aligned}$$
>
> is an isomorphism, so
>
> $$G/\operatorname{Ker}\phi \cong \operatorname{Im}\phi.$$

The First Isomorphism Theorem can sometimes be used to help identify a familiar, standard group isomorphic to a quotient group $G/\operatorname{Ker}\phi$, where ϕ is a homomorphism.

The isomorphism f in the First Isomorphism Theorem has domain the set of cosets of $\operatorname{Ker}\phi$ in $(G, \circ)$, codomain $\operatorname{Im}\phi$, and rule

coset $\longmapsto$ element of $\operatorname{Im}\phi$ that is the image under ϕ of each element of the coset.





The elements of the domain of f are whole cosets, not individual elements of $(G, \circ)$.

> **Corollary E56**
>
> Let $(G, \circ)$ be a finite group and let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism. Then
>
> $$|\operatorname{Ker}\phi| \times |\operatorname{Im}\phi| = |G|.$$

If $(G, \circ)$ and $(H, *)$ are finite groups, then the following numerical relationships hold for any homomorphism $\phi : (G, \circ) \longrightarrow (H, *)$:

$|\operatorname{Ker}(\phi)|$ divides $|G|$ (by Lagrange's Theorem)

$|\operatorname{Im}(\phi)|$ divides $|H|$ (by Lagrange's Theorem)

$|\operatorname{Im}(\phi)|$ divides $|G|$ (by Corollary E56).

In particular, the order of $\operatorname{Im}\phi$ is a common factor of the orders of the domain group $(G, \circ)$ and the codomain group $(H, *)$.

Unit E4 Group actions

1 Group actions

1. Group action definition

Let $(G, \circ)$ be a group with identity element e, and let X be a set. Suppose that for each element g in G and each element x in X an object $g \wedge x$ is defined in some way.

The effect $\wedge$ of $(G, \circ)$ on X is a **group action** (or simply an **action**) of $(G, \circ)$ on X if the following three **group action axioms** hold.

GA1 Closure For each $g \in G$ and each $x \in X$,

$$g \wedge x \in X.$$

GA2 Identity For each $x \in X$,

$$e \wedge x = x.$$

GA3 Composition For all $g, h \in G$ and all $x \in X$,

$$g \wedge (h \wedge x) = (g \circ h) \wedge x.$$

We also say that the group $(G, \circ)$ **acts on** X.

2. A **permutation** of a set X is a one-to-one and onto function from X to itself. (The set X may be either finite or infinite.)

Such a function **permutes** the elements of X.

Theorem E57

Let $\wedge$ be an action of a group G on a set X. Then $\wedge$ has the following properties.

(a) For each g in G, if x and y are elements of X such that $g \wedge x = g \wedge y$, then $x = y$.

(b) For each g in G, if y is an element of X then there is an element x of X such that $g \wedge x = y$.

Parts (a) and (b) of Theorem E57 say that the effect of each element of the group G on the elements of the set X is that of a one-to-one and onto function, respectively, from X to itself.

Thus in an action of a group G on a set X, each group element of G behaves as a permutation of the set X. However, two or more elements of G may behave like the same permutation.

A **faithful** group action is an action of a group G on a set X in which no two elements of G behave as the same permutation of X.

3. The **natural action** of a group G on a set X, or simply **the action** of G on X, is the obvious mapping effect of the elements of G on the elements of X, if there is one. For example, the group S_3 has a natural action on the set $\{1, 2, 3\}$, and the group $S(\square)$ has a natural action on the set of vertices of the square.

We also say that G **acts on** X **in the natural way**.

Proposition E58

For any natural number n, the usual mapping effect of the group S_n or any of its subgroups on the set $\{1, 2, \ldots, n\}$ of symbols is a group action.

4. Actions of groups of symmetries on sets of figures

A **group of symmetries** of a figure F is the symmetry group of F or one of its subgroups.

The **image** $g(A)$ of a figure A under an isometry g (such as a symmetry) is the set of points $\{g(P) : P \in A\}$.

Theorem E59

Let G be a group of symmetries of a figure F in $\mathbb{R}^2$, and let X be a set of figures in $\mathbb{R}^2$. Let $\wedge$ be defined by

$$g \wedge A = g(A),$$

for all $g \in G$ and all $A \in X$. Then $\wedge$ is a group action if and only if axiom GA1 (closure) holds.

The same is true if $\mathbb{R}^2$ is replaced by $\mathbb{R}^3$.

5. Actions of groups of symmetries on sets of coloured figures

A **coloured figure** is a figure whose points have each been assigned a colour from a finite set of colours.

The **image** $g(A)$ of a coloured figure A under an isometry g (such as a symmetry) is the coloured figure whose points are those in the set $\{g(P) : P \in A\}$, with each point $g(P)$ in $g(A)$ assigned the same colour as the corresponding point P in A.

Theorem E59 holds if 'figures' is replaced by 'coloured figures', as follows.

Theorem E60

Let G be a group of symmetries of a figure F in $\mathbb{R}^2$, and let X be a set of coloured figures in $\mathbb{R}^2$. Let $\wedge$ be defined by

$$g \wedge A = g(A),$$

for all $g \in G$ and all $A \in X$. Then $\wedge$ is a group action if and only if axiom GA1 (closure) holds.

The same is true if $\mathbb{R}^2$ is replaced by $\mathbb{R}^3$.

6. A group action arising from matrix multiplication

Theorem E61

Let $\wedge$ be defined by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \wedge (x, y) = (ax + by, cx + dy)$$

for all matrices $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(2)$ and all points $(x, y) \in \mathbb{R}^2$.

Then $\wedge$ is an action of the group $\mathrm{GL}(2)$ on the set $\mathbb{R}^2$.

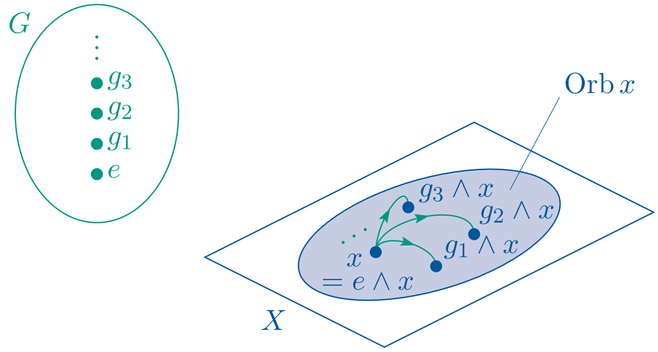
2 Orbits and stabilisers

7. Orbits

Let $\wedge$ be an action of a group G on a set X, and let x be an element of X. The **orbit** of x under $\wedge$, denoted by $\mathrm{Orb}\, x$, is

$$\mathrm{Orb}\, x = \{g \wedge x : g \in G\}.$$

That is, $\mathrm{Orb}\, x$ is the set of elements of X that can be obtained from x under the action of G. It is the subset of X that we obtain if we act on x using each element of G in turn.



We have: if $y = g \wedge x$, then $x = g^{-1} \wedge y$.

Theorem E62

Let $\wedge$ be an action of a group G on a set X. Then the distinct orbits of the elements of X under $\wedge$ form a partition of X.

The **orbits of a group action** are the distinct orbits of elements under the group action.

Strategy E7 Finding orbits

To find the orbits of an action of a group G on a finite set X, do the following.

1. Choose any element x of X, and find its orbit.
2. Choose any element of X not yet assigned to an orbit, and find its orbit.
3. Repeat step 2 until X is partitioned.

Strategy E7 can be used to find the orbits of an action of a group G on an *infinite* set X when there are only finitely many orbits.

8. Orbits of group actions on $\mathbb{R}^2$

The orbits of a group action can be described geometrically when the set on which the group acts is the plane $\mathbb{R}^2$. Such a group action may have only finitely many orbits, or infinitely many.

To find the orbits of a group action on $\mathbb{R}^2$ it is helpful to do the following.

1. Start by finding an expression for the orbit of a general point (x, y) in $\mathbb{R}^2$ under the group action.
2. Use this expression to find the orbits of a few particular points, to try to get an idea of how the plane might split up into orbits.
3. Use algebraic or geometric arguments to confirm what you think.

Keep in mind that the orbits partition the plane, so to find more orbits, you need to consider points that do not lie in the orbits that you have already found.

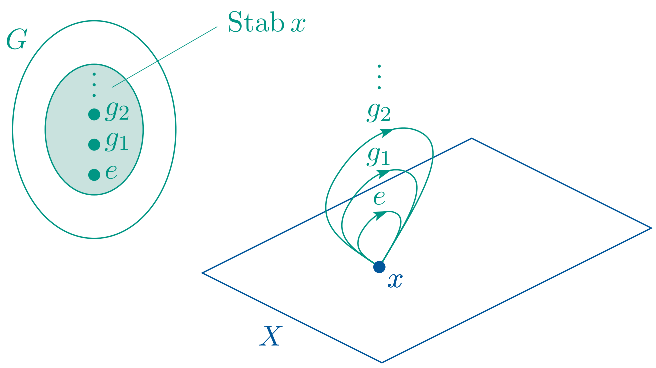
Orbits of group actions on $\mathbb{C}$ can be found in a similar way and also be described geometrically.

9. Stabilisers

Let $\wedge$ be an action of a group G on a set X, and let x be an element of X. The **stabiliser** of x under $\wedge$, denoted by $\operatorname{Stab} x$, is given by

$$\operatorname{Stab} x = \{g \in G : g \wedge x = x\}.$$

That is, $\operatorname{Stab} x$ is the set of elements of G that fix x (map x to itself).



Theorem E63

Let $\wedge$ be an action of a group G on a set X. Then, for each element x of X, the set $\operatorname{Stab} x$ is a subgroup of G.

3 The Orbit–Stabiliser Theorem

10. For actions of finite groups we have the following results.

Theorem E64 Orbit–Stabiliser Theorem

Suppose that the finite group G acts on the set X. Then, for each element x in X,

$$|\operatorname{Orb} x| \times |\operatorname{Stab} x| = |G|.$$

Corollary E65

Suppose that the finite group G acts on the set X. Then, for each element x in X, the number of elements in $\operatorname{Orb} x$ divides the order of G.

11. Left cosets of stabilisers

The following theorem and corollary are used to prove the Orbit–Stabiliser Theorem.

Theorem E66

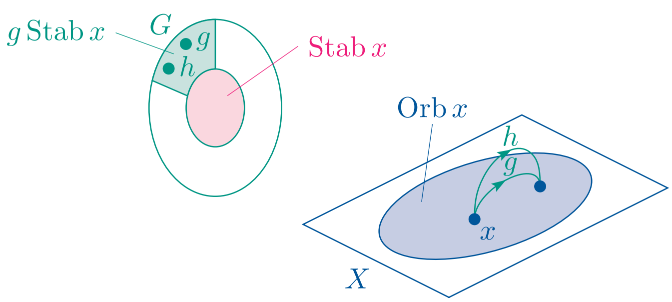
Let $\wedge$ be an action of a group G on a set X, let x be an element of X and let g and h be elements of G. Then

$$g \wedge x = h \wedge x$$

if and only if

g and h lie in the same left coset of $\operatorname{Stab} x$.

Thus, if G is a group acting on a set X and x is an element of X, then the left cosets of $\operatorname{Stab} x$ in G are the sets of elements of G that map x to the same element of X.

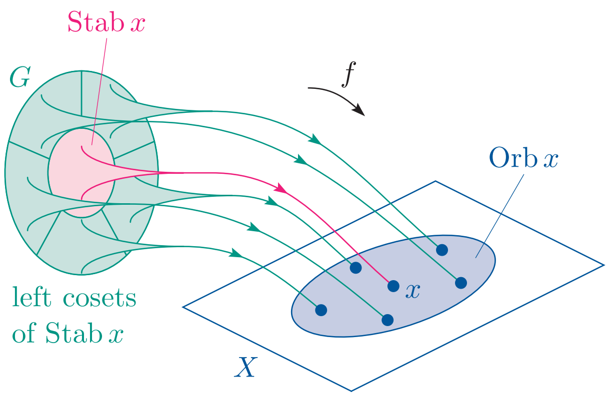


Corollary E67

Let $\wedge$ be an action of a group G on a set X and let x be an element of X. Then the mapping f given by

$$\begin{aligned} f: \text{set of left cosets of } \operatorname{Stab} x &\longrightarrow \operatorname{Orb} x \\ g \operatorname{Stab} x &\longmapsto g \wedge x \end{aligned}$$

is one-to-one and onto.



12. Groups acting on groups: conjugation

Proposition E68

Let G be a group, and let $\wedge$ be defined by

$$g \wedge x = gxg^{-1}$$

for all $g, x \in G$. Then $\wedge$ is an action of G on itself.

Thus conjugation is an action of a group on itself. Its orbits are the conjugacy classes of G, so a special case of Corollary E65 is Theorem E27: the number of elements in a conjugacy class divides the order of the group.

13. Groups acting on groups: Lagrange's Theorem

Let H be a subgroup of a group G, and let $\wedge$ be defined by

$$h \wedge g = hg$$

for all $h \in H$ and $g \in G$. Then $\wedge$ is an action of H on G.

Its orbits are the right cosets of H in G.

Thus another special case of Corollary E65 is Lagrange's Theorem: if H is a subgroup of a finite group G, then the order of H divides the order of G.

14. Groups acting on groups: homomorphisms

Let $\phi : (G, \circ) \longrightarrow (H, *)$ be a homomorphism, and let $\wedge$ be defined by

$$g \wedge h = \phi(g) * h,$$

for all $g \in G$ and $h \in H$. Then $\wedge$ is an action of the group $(G, \circ)$ on the group $(H, *)$.

If G is finite, then applying the Orbit–Stabiliser Theorem to this group action with $x = e_H$ gives

$$|\operatorname{Im} \phi| \times |\operatorname{Ker} \phi| = |G|,$$

which is the conclusion of Corollary E56.

Thus Corollary E56 is a special case of the Orbit–Stabiliser Theorem.

4 The Counting Theorem

15. Multiplication Principle

If we have k successive choices to make, and the ith choice involves choosing from n_i options, for each $i = 1, 2, \ldots, k$, then the total number of ways to make all k choices is

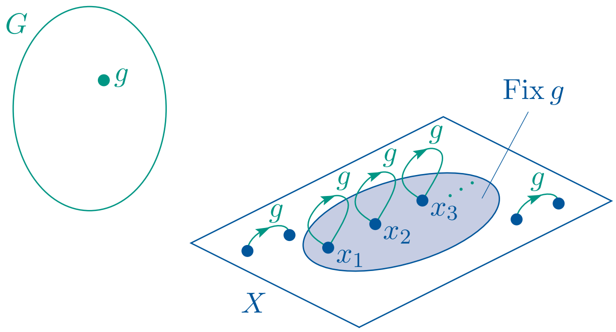
$$n_1 \times n_2 \times \cdots \times n_k.$$

16. Fixed sets

Let $\wedge$ be an action of a group G on a set X, and let g be an element of G. The **fixed set** of g under $\wedge$, denoted by $\operatorname{Fix} g$, is given by

$$\operatorname{Fix} g = \{x \in X : g \wedge x = x\}.$$

That is, $\operatorname{Fix} g$ is the set of elements of X that are fixed by g.



It is an element of the *group* G, not an element of the *set* X, that has a fixed set.

- The fixed set of an element g in G is the set of all elements of X that are fixed by g.
- The stabiliser of an element x in X is the set of all elements of G that fix x.

$\operatorname{Fix} g$ is a subset of X, whereas $\operatorname{Stab} x$ is a subgroup of G.

The fixed point set of a symmetry g of a figure F is the set of all points of F that are fixed by g (see **17.** on page 99). This is a special case of a fixed set; it is the fixed set of g under the natural action of the symmetry group $S(F)$ on the set of points in F.

17. Permutation method for finding $|\operatorname{Fix} g|$

Let F be a figure with a finite number of parts, each of which is to be coloured with one of c colours, and let X be the set of all coloured figures that can be obtained by colouring F in this way. Let G be a group of symmetries of F that acts on X, and let g be a symmetry in G. To find the size of $\operatorname{Fix} g$, do the following.

1. Label each of the parts of F to be coloured, using a different symbol for each of them.
2. Hence express the effect of g on these parts as a permutation in cycle form, *including any 1-cycles*.
3. Then $|\operatorname{Fix} g| = c^k$ where k is the number of cycles of the permutation, *including any 1-cycles*.

18. The Counting Theorem

Theorem E69 Counting Theorem

Let $\wedge$ be an action of a finite group G on a finite set X. Then the number of orbits of $\wedge$ is given by

$$\frac{1}{|G|} \sum_{g \in G} |\operatorname{Fix} g|.$$

That is, to find the number of orbits of an action of a finite group G on a finite set X, determine the number $|\operatorname{Fix} g|$ for each element g in G, add up all these numbers, and divide the total by the order of G.

19. A **counting problem** is a problem that asks how many objects there are of a particular type.

A counting problem involving symmetry can often be solved by interpreting it as a problem of finding the number of orbits of the action of a group of symmetries on a set of coloured figures. We can use the Counting Theorem to determine the number of orbits.

The working often can be simplified by noting that symmetries of the same geometric type (that is, symmetries that are conjugate in $S(F)$) fix the same number of coloured figures.

Book F Analysis 2

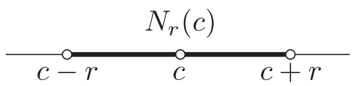
Unit F1 Limits

1 Limits of functions

1. A **punctured neighbourhood** of a point c is a bounded open interval with midpoint c, from which the point c itself has been removed:

$$N_r(c) = (c - r, c) \cup (c, c + r), \quad \text{where } r > 0,$$

is a punctured neighbourhood of length $2r$ with centre c.



2. Sequential definition of a limit

Let f be a function defined on a punctured neighbourhood $N_r(c)$ of c. Then $f(x)$ **tends to the limit l as x tends to c** if $l \in \mathbb{R}$ and

for each sequence (x_n) in $N_r(c)$ such that $x_n \to c$,

$f(x_n) \to l$.

We write

$$\lim_{x \to c} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c.$$

Strategy F1 Showing a limit does not exist

Let f be a real function defined on a punctured neighbourhood $N_r(c)$ of c.

To show that $\lim_{x \to c} f(x)$ does not exist, either:

- find two sequences (x_n) and (y_n) in $N_r(c)$ which tend to c, such that $(f(x_n))$ and $(f(y_n))$ have different limits, or
- find a sequence (x_n) in $N_r(c)$ which tends to c such that $f(x_n) \to \infty$ or $f(x_n) \to -\infty$.

Theorem F1

$$\lim_{x \to 0} \frac{\sin x}{x} = 1.$$

3. Limits and continuity

Theorem F2

Let f be a function defined on an open interval I, with $c \in I$. Then

f is continuous at c

if and only if

$$\lim_{x \to c} f(x) = f(c).$$

4. Rules for limits

Theorem F3 Combination Rules

If $\lim_{x \to c} f(x) = l$ and $\lim_{x \to c} g(x) = m$, then:

Sum Rule $\lim_{x \to c} (f(x) + g(x)) = l + m$

Multiple Rule $\lim_{x \to c} \lambda f(x) = \lambda l$, for $\lambda \in \mathbb{R}$

Product Rule $\lim_{x \to c} f(x)g(x) = lm$

Quotient Rule $\lim_{x \to c} f(x)/g(x) = l/m$, provided that $m \neq 0$.

Theorem F4 Composition Rule

If $\lim_{x \to c} f(x) = l$ and $\lim_{x \to l} g(x) = L$, then

$$\lim_{x \to c} g(f(x)) = L,$$

provided that

either $f(x) \neq l$, for all x in some $N_r(c)$, where $r > 0$,

or g is defined at l and continuous at l.

Theorem F5 Squeeze Rule

Let f, g and h be functions defined on $N_r(c)$, for some $r > 0$. If

(a) $g(x) \leq f(x) \leq h(x)$, for $x \in N_r(c)$

(b) $\lim_{x \to c} g(x) = \lim_{x \to c} h(x) = l$,

then $\lim_{x \to c} f(x) = l$.

5. Evaluating the limit of a composition

Strategy F2 Evaluating a limit - Composition Rule

To use the Composition Rule to evaluate a limit of a function of the form $g(f(x))$ as $x \to c$, do the following.

1. Substitute $u = f(x)$ and show that, for some l,
$$u = f(x) \to l \text{ as } x \to c.$$
2. Show that, for some L,
$$g(u) \to L \text{ as } u \to l.$$
3. Check that one of the provisos holds.
4. Deduce that
$$g(f(x)) \to L \text{ as } x \to c.$$

***Warning*:** Be careful not to omit step 3 from Strategy F2.

6. Basic limits

Theorem F6 Three basic limits

- $\lim_{x \to 0} \frac{\sin x}{x} = 1$
- $\lim_{x \to 0} \frac{1 - \cos x}{x} = 0$
- $\lim_{x \to 0} \frac{e^x - 1}{x} = 1$

7. One-sided limits

Let f be a function defined on $(c, c+r)$, for some $r > 0$. Then $f(x)$ **tends to the limit l as x tends to c from the right** if

for each sequence (x_n) in $(c, c+r)$ such that $x_n \to c$,
$$f(x_n) \to l.$$

We write

$$\lim_{x \to c^+} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c^+.$$

A limit **as x tends to c from the left** is defined similarly, with $(c, c+r)$ replaced by $(c-r, c)$. We write

$$\lim_{x \to c^-} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c^-.$$

We refer to

$$\lim_{x \to c^+} f(x) \quad \text{and} \quad \lim_{x \to c^-} f(x)$$

as **right** and **left limits**, respectively.

Sometimes both right and left limits exist but are different.

8. Results for one-sided limits

Theorem F7

Let the function f be defined on $N_r(c)$, for some $r > 0$. Then
$$\lim_{x \to c} f(x) = l$$
if and only if
$$\lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x) = l.$$

Versions of the Combination, Composition and Squeeze Rules, and Strategy F2, can also be used to determine one-sided limits: here $\lim_{x \to c}$ and $N_r(c)$ are replaced by $\lim_{x \to c^+}$ and $(c, c+r)$, or $\lim_{x \to c^-}$ and $(c-r, c)$, respectively.

A one-sided limit can be shown not to exist using a version of Strategy F1 where the sequences (x_n) and (y_n) are chosen to tend to c from the right, or left, as appropriate.

Theorem F8 is a version of Theorem F2 for one-sided limits. There is an analogous version for left limits.

Theorem F8

Let f be a function whose domain is an interval I with a finite left-hand endpoint c that lies in I. Then
$$f \text{ is continuous at } c$$
if and only if
$$\lim_{x \to c^+} f(x) = f(c).$$

2 Asymptotic behaviour of functions

9. Functions tending to plus/minus infinity

Let the function f be defined on $N_r(c)$, for some $r > 0$. Then $f(x)$ **tends to ∞ as x tends to c** if

for each sequence (x_n) in $N_r(c)$ such that $x_n \to c$,
$f(x_n) \to \infty$.

We write

$$f(x) \to \infty \quad \text{as } x \to c.$$

The statements

$$f(x) \to -\infty \quad \text{as } x \to c,$$
$$f(x) \to \infty \text{ (or } -\infty) \quad \text{as } x \to c^+ \text{ (or } c^-),$$

are defined in a similar way, with ∞ replaced by $-\infty$ and $N_r(c)$ replaced by the open interval $(c, c+r)$ or $(c-r, c)$, where $r > 0$, as appropriate.

***Warning*:** Do not use the notation

$$\lim_{x\to c} f(x) = \infty$$

as this can give the misleading impression that infinity can be treated in a similar way to a finite limit.

10. Rules for functions tending to infinity

Theorem F9 Reciprocal Rule

If the function f satisfies the conditions

1. $f(x) > 0$ for $x \in N_r(c)$, for some $r > 0$
2. $f(x) \to 0$ as $x \to c$,

then

$$\frac{1}{f(x)} \to \infty \quad \text{as } x \to c.$$

The Reciprocal Rule can be applied with $x \to c$ replaced by $x \to c^+$, or $x \to c^-$, and $N_r(c)$ replaced by $(c, c+r)$ and $(c-r, c)$, as appropriate.

There are also versions of the Combination Rules and the Squeeze Rule for functions which tend to ∞ (or $-\infty$) as x tends to c, c^+ or c^-, for example:

Theorem F10 Combination Rules

If $f(x) \to \infty$ as $x \to c$ and $g(x) \to \infty$ as $x \to c$, then:

Sum Rule $f(x) + g(x) \to \infty$ as $x \to c$

Multiple Rule $\lambda f(x) \to \infty$ as $x \to c$, for $\lambda \in \mathbb{R}^+$

Product Rule $f(x)g(x) \to \infty$ as $x \to c$.

11. Let the function f be defined on (R, ∞), for some real number R. Then $f(x)$ **tends to l as x tends to ∞** if

for each sequence (x_n) in (R, ∞) such that $x_n \to \infty$,
$f(x_n) \to l$.

In this case, we write

$$f(x) \to l \quad \text{as } x \to \infty.$$

The statement

$$f(x) \to l \quad \text{as } x \to -\infty$$

is defined similarly, with ∞ replaced by $-\infty$, and (R, ∞) replaced by $(-\infty, R)$.

The letter l denotes either a real number or one of the symbols ∞ or $-\infty$.

When l is a real number, we also use the notations

$$\lim_{x\to\infty} f(x) = l \quad \text{and} \quad \lim_{x\to-\infty} f(x) = l.$$

Theorem F11 Basic asymptotic behaviour

If $n \in \mathbb{N}$, then

- $x^n \to \infty$ as $x \to \infty$
- $\dfrac{1}{x^n} \to 0$ as $x \to \infty$.

12. The **dominant term** of a quotient involving the real variable x is the term in x (without its coefficient) which eventually has the largest absolute value.

This is similar to determining the behaviour of sequences defined by quotients using Strategy D7 (see **13.** on page 80).

13. Rules for functions as $x \to \infty$

Versions of the Reciprocal Rule and the Combination Rules can be used to obtain results about the behaviour of functions as $x \to \infty$ or $-\infty$: here c and $N_r(c)$ are replaced by ∞ and (R, ∞), or $-\infty$ and $(-\infty, R)$, respectively.

Theorem F12 Squeeze Rule

Let f, g and h be functions defined on some interval (R, ∞).

(a) If f, g and h satisfy the conditions

1. $g(x) \leq f(x) \leq h(x)$, for $x \in (R, \infty)$
2. $\lim\limits_{x\to\infty} g(x) = \lim\limits_{x\to\infty} h(x) = l$

where l is a real number, then

$$\lim_{x\to\infty} f(x) = l.$$

(b) If f and g satisfy the conditions

1. $f(x) \geq g(x)$, for $x \in (R, \infty)$
2. $g(x) \to \infty$ as $x \to \infty$

then

$$f(x) \to \infty \quad \text{as } x \to \infty.$$

14. Asymptotic behaviour: standard results

Theorem F13

(a) If $a_0, a_1, \ldots, a_{n-1} \in \mathbb{R}$, where $n \in \mathbb{N}$, and

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$

then

- $p(x) \to \infty \quad \text{as } x \to \infty$
- $\dfrac{1}{p(x)} \to 0 \quad \text{as } x \to \infty.$

(b) For each $n = 0, 1, 2, \ldots$, we have

- $\dfrac{e^x}{x^n} \to \infty \quad \text{as } x \to \infty$
- $\dfrac{x^n}{e^x} \to 0 \quad \text{as } x \to \infty.$

(c) We have

$$\log x \to \infty \quad \text{as } x \to \infty,$$

but, for each constant $a > 0$, we have

$$\frac{\log x}{x^a} \to 0 \quad \text{as } x \to \infty.$$

15. Composition Rule for asymptotic behaviour

Versions of the Composition Rule and Strategy F2 (see **5.** on page 110) can be used to deduce the asymptotic behaviour of composites of functions which have many types of asymptotic behaviour: the letters l and L denote either a real number or one of the symbols ∞ or $-\infty$.

Notice that if l is either ∞ or $-\infty$, then the first proviso of the Composition Rule is automatically satisfied.

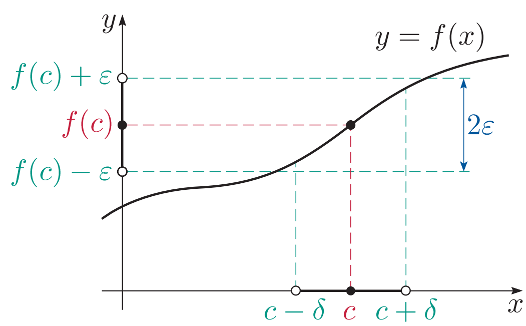
3 Continuity: the classical definition

16. ε-δ definition of continuity

Let the function f have domain A and let $c \in A$. Then f is **continuous** at c if

for each $\varepsilon > 0$, there exists $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x \in A \text{ with } |x - c| < \delta.$$



Theorem F14

The ε-δ definition and the sequential definition of continuity are equivalent.

On the whole it is usually easier to use the sequential definition to prove

- discontinuity
- continuity of simpler functions.

The ε-δ definition can work better for proving continuity of more complicated functions.

Strategy F3 Showing continuity using ε-δ for polynomial functions

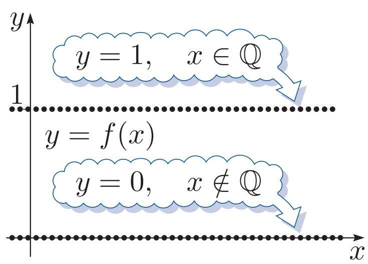
To use the ε-δ definition to prove that a polynomial function f with domain A is continuous at a point $c \in A$, let $\varepsilon > 0$ be given and carry out the following steps.

1. Use algebraic manipulation to express the difference $f(x) - f(c)$ as a product of the form $(x - c)g(x)$.
2. Obtain an upper bound of the form $|g(x)| \leq M$, for $|x - c| \leq r$, where $r > 0$ is chosen so that $[c - r, c + r] \subset A$. (The Triangle Inequality is often useful here.)
3. Use the fact that $|f(x) - f(c)| \leq M|x - c|$, for $|x - c| \leq r$, to choose $\delta > 0$ such that

$$|f(x) - f(c)| < \varepsilon, \quad \text{for all } x \in A \text{ with } |x - c| < \delta.$$

17. The **Dirichlet function** has domain $\mathbb{R}$ and rule

$$f(x) = \begin{cases} 1, & \text{if } x \text{ is rational,} \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$



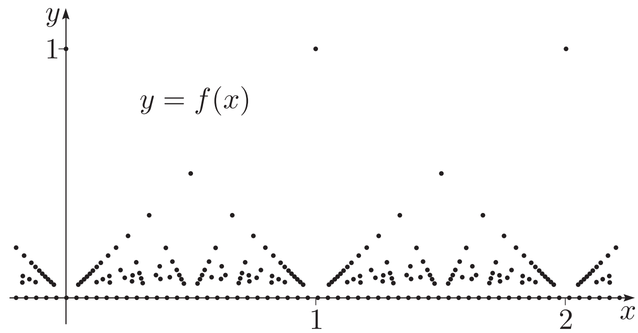
Theorem F15

The Dirichlet function is discontinuous at every point of $\mathbb{R}$.

18. The **Riemann function** has domain $\mathbb{R}$ and rule

$$f(x) = \begin{cases} 1/q, & \text{if } x \text{ is a rational } p/q, (q > 0), \\ 0, & \text{if } x \text{ is irrational.} \end{cases}$$

Note, p/q is expressed in lowest terms.



Theorem F16

The Riemann function is discontinuous at each rational point of $\mathbb{R}$ and continuous at each irrational point.

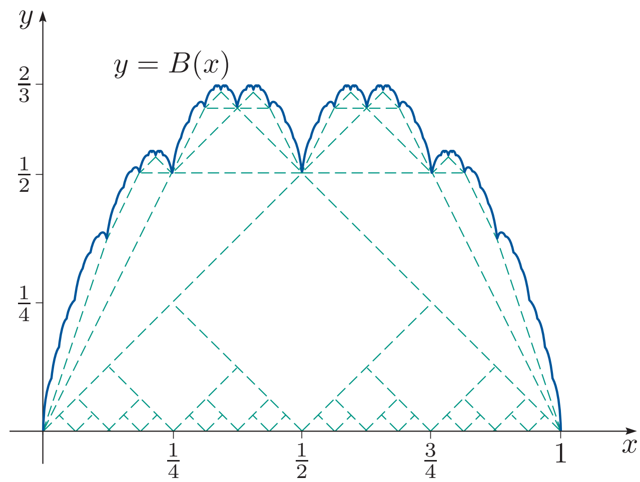
19. The **sawtooth function** has domain $\mathbb{R}$ and rule

$$s(x) = \begin{cases} x - \lfloor x \rfloor, & \text{if } 0 \leq x - \lfloor x \rfloor \leq \frac{1}{2}, \\ 1 - (x - \lfloor x \rfloor), & \text{if } \frac{1}{2} < x - \lfloor x \rfloor < 1, \end{cases}$$

where $\lfloor x \rfloor$ is the integer part function.

The **blancmange function** B is obtained by forming an infinite series of functions related to s:

$$B(x) = s(x) + \tfrac{1}{2}s(2x) + \tfrac{1}{4}s(4x) + \tfrac{1}{8}s(8x) + \cdots$$
$$= \sum_{n=0}^{\infty} \frac{1}{2^n} s(2^n x).$$



Theorem F17

The blancmange function is continuous.

20. **ε-δ definition of a limit**

Let f be a function defined on a punctured neighbourhood $N_r(c)$ of c. Then $f(x)$ **tends to the limit l as x tends to c** if

$$\text{for each } \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that}$$
$$|f(x) - l| < \varepsilon, \quad \text{for all } x \text{ with } 0 < |x - c| < \delta.$$

We write

$$\lim_{x \to c} f(x) = l \quad \text{or} \quad f(x) \to l \text{ as } x \to c.$$

The ε-δ definition and the sequential definition of a limit are equivalent.

4 Uniform continuity

21. An **interior point** c of an interval I is a point that is not an endpoint of I.

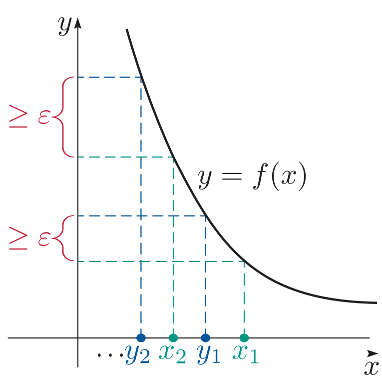
22. A function f defined on an interval I is **uniformly continuous** on I if

$$\text{for each } \varepsilon > 0, \text{ there exists } \delta > 0 \text{ such that}$$
$$|f(x) - f(y)| < \varepsilon, \quad \text{for all } x, y \in I \text{ with } |x - y| < \delta.$$

Theorem F18

Let the function f be defined on an interval I. Then f is not uniformly continuous on I if and only if there exist two sequences (x_n) and (y_n) in I, and $\varepsilon > 0$, such that

1. $|x_n - y_n| \to 0$ as $n \to \infty$
2. $|f(x_n) - f(y_n)| \geq \varepsilon$, for $n = 1, 2, \ldots$.



Strategy F4 Showing uniform continuity/discontinuity

- To prove that a function f is uniformly continuous on an interval I, find an expression for $\delta > 0$ in terms of a given $\varepsilon > 0$ such that
$$|f(x) - f(y)| < \varepsilon, \quad \text{for all } x, y \in I \text{ with } |x - y| < \delta.$$
- To prove that a function f is *not* uniformly continuous on an interval I, find two sequences (x_n) and (y_n) in I, and $\varepsilon > 0$, such that
$$|x_n - y_n| \to 0 \quad \text{as } n \to \infty,$$
and
$$|f(x_n) - f(y_n)| \geq \varepsilon, \quad \text{for } n = 1, 2, \ldots.$$

***Alternative*:** If it is applicable, it is easier to apply Theorem F19 (below) to show that a function is uniformly continuous.

When using Strategy F4 to prove that a function is not uniformly continuous, aim to choose the terms x_n and y_n close together at points of I where the graph of f is steep.

Theorem F19

If the function f is continuous on a bounded closed interval $[a, b]$, then f is uniformly continuous on $[a, b]$.

Theorem F20 Bolzano–Weierstrass Theorem

Any bounded sequence has a convergent subsequence.

Unit F2 Differentiation

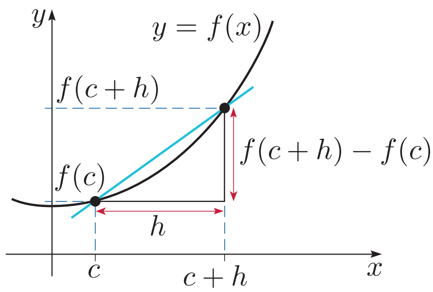
1 Differentiable functions

1. Let f be defined on an open interval I, and let $c \in I$.

The **difference quotient** for f at c is

$$\frac{f(x) - f(c)}{x - c}, \quad \text{or} \quad Q(h) = \frac{f(c+h) - f(c)}{h},$$

where $x \neq c$, $h \neq 0$.



The **gradient** (or slope), of the graph of f at the point $(c, f(c))$ is

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c}, \quad \text{or} \quad \lim_{h \to 0} Q(h),$$

provided that the limit exists.

2. Let f be defined on an open interval I, and let $c \in I$.

The **derivative** of f at c is

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c},$$

that is,

$$\lim_{h \to 0} Q(h),$$

provided that this limit exists.

The function f is

- **differentiable at c** if this limit exists
- **differentiable** (on its domain) if f is differentiable at each point of its domain.

The derivative of f at c is denoted by $f'(c)$ and the function $f' : x \longmapsto f'(x)$ is called the **derivative**, or the **derived function**, of f.

Differentiation is the operation of obtaining $f'(x)$ from $f(x)$.

Strategy F5 Showing a function is not differentiable

To prove that a function is not differentiable at a point, show that $\lim_{h \to 0} Q(h)$ does not exist by doing either of the following.

- Find two null sequences (h_n) and (k_n) with non-zero terms such that the sequences $(Q(h_n))$ and $(Q(k_n))$ have different limits.
- Find a null sequence (h_n) with non-zero terms such that $Q(h_n) \to \infty$ or $Q(h_n) \to -\infty$.

Theorem F21 Basic derivatives

- If $f(x) = k$, where $k \in \mathbb{R}$, then $f'(x) = 0$.
- If $f(x) = x^n$, where $n \in \mathbb{N}$, then $f'(x) = nx^{n-1}$.
- If $f(x) = \sin x$, then $f'(x) = \cos x$.
- If $f(x) = \cos x$, then $f'(x) = -\sin x$.
- If $f(x) = e^x$, then $f'(x) = e^x$.

3. Let f be differentiable on an open interval I, and let $c \in I$.

The function f is **twice differentiable at c**, if the derivative f' is differentiable at c. The **second derivative of f at c** is the number $f''(c) = (f')'(c)$.

The **second derivative** (or **second derived function**) of f is the function f'', also denoted by $f^{(2)}$.

Higher-order derivatives of f, denoted by $f^{(3)} = f'''$, $f^{(4)}$, and so on, are defined similarly.

Some functions (for example, all those in Theorem F21) can be differentiated as many times as we like. However, not every derivative is differentiable at all points of its domain (for example, the derivative of the function f, given by $f(x) = x^2$ for $x > 0$ and $f(x) = 0$ for $x \leq 0$, is not differentiable at $x = 0$).

4. In Leibniz notation $f'(x)$ is written as $\dfrac{dy}{dx}$, and $f''(x)$ as $\dfrac{d^2y}{dx^2}$, where $y = f(x)$.

5. Let f be defined on an interval I, and let $c \in I$.

The **left derivative** of f at c is

$$f'_L(c) = \lim_{x \to c^-} \frac{f(x) - f(c)}{x - c} = \lim_{h \to 0^-} Q(h),$$

provided that this limit exists. If the limit exists, we say that f is **left differentiable** at c.

Similarly, the **right derivative** of f at c is

$$f'_R(c) = \lim_{x \to c^+} \frac{f(x) - f(c)}{x - c} = \lim_{h \to 0^+} Q(h),$$

provided that this limit exists. If the limit exists, we say that f is **right differentiable** at c.

Theorem F22

Let f be defined on an open interval I, and let $c \in I$.

(a) If f is differentiable at c, then f is both left differentiable and right differentiable at c, and

$$f'_L(c) = f'_R(c) = f'(c).$$

(b) If f is both left differentiable and right differentiable at c, and $f'_L(c) = f'_R(c)$, then f is differentiable at c and

$$f'(c) = f'_L(c) = f'_R(c).$$

6. Glue Rule for differentiable functions

Theorem F23 Glue Rule

Let f be defined on an open interval I, and let $c \in I$. If there are functions g and h defined on I such that

1. $f(x) = g(x)$, for $x \in I$, $x < c$,
   $f(x) = h(x)$, for $x \in I$, $x > c$,
2. $f(c) = g(c) = h(c)$, and
3. g and h are differentiable at c,

then f is differentiable at c if and only if $g'(c) = h'(c)$. If f is differentiable at c, then $f'(c) = g'(c) = h'(c)$.

7. Differentiability is a local property because it depends on the behaviour of the function in any open interval (no matter how short).

The restriction of a differentiable function to an open subinterval of its domain is a differentiable function.

8. Continuity and differentiability

Let f be defined on an open interval I, and let $c \in I$.

Theorem F24 If f is differentiable at c, then f is continuous at c.

Corollary F25 If f is discontinuous at c, then f is not differentiable at c.

***Warning*:** A function can be continuous at a point and not differentiable at that point (for example, the blancmange function is continuous everywhere, but not differentiable anywhere).

Theorem F26

The blancmange function is not differentiable at any point $c \in \mathbb{R}$.

2 Rules for differentiation

9. Rules for differentiable functions

Theorem F30 Combination Rules

Let f and g be defined on an open interval I, and let $c \in I$. If f and g are differentiable at c, then so are the following functions.

Sum Rule $f + g$, with derivative

$$(f + g)'(c) = f'(c) + g'(c)$$

Multiple Rule λf, with derivative

$$(\lambda f)'(c) = \lambda f'(c) \quad (\lambda \in \mathbb{R})$$

Product Rule fg, with derivative

$$(fg)'(c) = f'(c)g(c) + f(c)g'(c)$$

Quotient Rule f/g, provided that $g(c) \neq 0$, with derivative

$$\left(\frac{f}{g}\right)'(c) = \frac{g(c)f'(c) - f(c)g'(c)}{(g(c))^2}.$$

Corollary F31

Let

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n,$$

for $x \in \mathbb{R}$, where $a_0, a_1, \ldots, a_n \in \mathbb{R}$. Then p is differentiable on $\mathbb{R}$, with derivative

$$p'(x) = a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

Theorem F32 Composition Rule (the Chain Rule)

Let f be defined on an open interval I, let g be defined on an open interval J such that $f(I) \subseteq J$, and let $c \in I$.

If f is differentiable at c and g is differentiable at $f(c)$, then $g \circ f$ is differentiable at c and

$$(g \circ f)'(c) = g'(f(c))f'(c).$$

Theorem F33 Inverse Function Rule

Let f be a function whose domain is an open interval I on which f is continuous and strictly monotonic. Then f has an inverse function f^{-1} with domain $J = f(I)$.

If f is differentiable on I and $f'(x) \neq 0$ for $x \in I$, then f^{-1} is differentiable on J. Also, if $c \in I$ and $d = f(c)$, then

$$(f^{-1})'(d) = \frac{1}{f'(c)}.$$

In Leibniz notation, if $u = f(x)$ and $y = g(u) = g(f(x))$, then the Composition Rule becomes

$$\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx},$$

and if $y = f(x)$ and $x = f^{-1}(y)$, then the Inverse Function Rule becomes

$$\frac{dx}{dy} = \frac{1}{dy/dx}.$$

10. The following functions are differentiable:

- polynomial and rational functions
- $f(x) = x^0$, on $\mathbb{R} - \{0\}$
- the exponential function
- the trigonometric and hyperbolic functions.

3 Rolle's Theorem

11. Let f be defined on an interval $[a, b]$. Then: $f(d)$ is the **maximum** of f on $[a, b]$ if $d \in [a, b]$ and

$$f(x) \leq f(d) \quad \text{for } x \in [a, b];$$

$f(c)$ is the **minimum** of f on $[a, b]$ if $c \in [a, b]$ and

$$f(x) \geq f(c) \quad \text{for } x \in [a, b].$$

An **extreme value** of f is a maximum or a minimum.

12. The function f has:

- a **local maximum** $f(c)$ at c if there is an open interval $I = (c - r, c + r)$, where $r > 0$, in the domain of f such that
  $$f(x) \leq f(c), \quad \text{for } x \in I$$
- a **local minimum** $f(c)$ at c if there is an open interval $I = (c - r, c + r)$, where $r > 0$, in the domain of f such that
  $$f(x) \geq f(c), \quad \text{for } x \in I$$
- a **local extreme value** $f(c)$ at c if $f(c)$ is either a local maximum or a local minimum.

13. A **stationary point** of f is a point c such that $f'(c) = 0$; we also say that f' **vanishes** at c.

Theorem F34 Local Extreme Value Theorem

If f has a local extreme value at c and f is differentiable at c, then

$$f'(c) = 0.$$

Corollary F35

Let f be continuous on the closed interval $[a, b]$ and differentiable on (a, b). Then the extreme values of f on $[a, b]$ can occur only at a or b, or at points x in (a, b) where $f'(x) = 0$.

14. **Extreme values and stationary points**

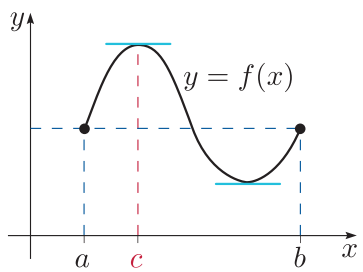
Strategy F6 Finding maxima/minima

To find the maximum and minimum of a function f that is continuous on $[a, b]$ and differentiable on (a, b), do the following.

1. Determine the points $c_1, c_2, \dots$ in (a, b) where f' is zero.
2. Hence determine the values of
$$f(a),\ f(b),\ f(c_1),\ f(c_2),\ \dots;$$
the greatest of these is the maximum and the least is the minimum.

Theorem F36 Rolle's Theorem

Let f be continuous on the closed interval $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$, then there exists a point c, with $a < c < b$, such that
$$f'(c) = 0.$$



4 Mean Value Theorem

15. The Mean Value Theorem generalises Rolle's Theorem: there is always a point where the tangent to the graph is parallel to the chord joining the endpoints.

Theorem F37 Mean Value Theorem

Let f be continuous on the closed interval $[a, b]$ and differentiable on (a, b). Then there exists a point c in (a, b) such that
$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

16. The **interior** of an interval I is the largest open subinterval of I (obtained by removing any endpoints of I).

Let f be continuous on an interval I and differentiable on the interior of I.

Theorem F38 Increasing–Decreasing Theorem

(a) If $f'(x) \geq 0$ for x in the interior of I, then f is increasing on I.

(b) If $f'(x) \leq 0$ for x in the interior of I, then f is decreasing on I.

Corollary F39 Zero Derivative Theorem

If $f'(x) = 0$ for all x in the interior of I, then f is constant on I.

Theorem F40 Second Derivative Test

Let f be a twice-differentiable function defined on an open interval I containing a point c such that $f'(c) = 0$ and f'' is continuous at c.

(a) If $f''(c) > 0$, then $f(c)$ is a local minimum of f.

(b) If $f''(c) < 0$, then $f(c)$ is a local maximum of f.

Strategy F7 Proving inequalities - Increasing–Decreasing

To prove that $g(x) \geq h(x)$, for $x \in [a, b]$, carry out the following steps.

1. Let
$$f(x) = g(x) - h(x),$$
and show that f is continuous on $[a, b]$ and differentiable on (a, b).
2. Prove that

either $f(a) \geq 0$ and $f'(x) \geq 0$ for $x \in (a, b)$,

or $f(b) \geq 0$ and $f'(x) \leq 0$ for $x \in (a, b)$.

There is a corresponding version of this strategy and the Increasing–Decreasing Theorem with the weak inequalities replaced by strict inequalities.

5 L'Hôpital's Rule

17. Cauchy's generalisation of the Mean Value Theorem involves two functions which can be interpreted as a pair of parametric equations describing the curve.

Theorem F41 Cauchy's Mean Value Theorem

Let f and g be continuous on $[a,b]$ and differentiable on (a,b). Then there exists a point $c \in (a,b)$ such that

$$f'(c)(g(b)-g(a)) = g'(c)(f(b)-f(a));$$

in particular, if $g(b) \neq g(a)$ and $g'(c) \neq 0$, then

$$\frac{f'(c)}{g'(c)} = \frac{f(b)-f(a)}{g(b)-g(a)}.$$

18. L'Hôpital's Rule

L'Hôpital's Rule enables us to evaluate limits of quotients where the Quotient Rule does not apply: limits of the form

$$\lim_{x\to c}\frac{f(x)}{g(x)},$$

where f and g are differentiable functions with $f(c) = g(c) = 0$.

Theorem F42 L'Hôpital's Rule

Let f and g be differentiable on an open interval I containing the point c, and suppose that $f(c) = g(c) = 0$. Then

$$\lim_{x\to c}\frac{f(x)}{g(x)} \text{ exists and equals } \lim_{x\to c}\frac{f'(x)}{g'(x)},$$

provided that the latter limit exists.

***Warning*:** Do not forget to check that all the conditions hold.

If $f'(c) = g'(c) = 0$, then a second application of L'Hôpital's Rule is needed.

Unit F3 Integration

1 The Riemann integral

1. Let the function f be defined on the closed interval $[a,b]$. Then the following hold on $[a,b]$.

- f is **bounded below** on $[a,b]$ with m as a **lower bound** if
  $$f(x) \geq m, \quad \text{for all } x \in [a,b].$$
- m is the **minimum** of f on $[a,b]$ if
  1. m is a lower bound for f on $[a,b]$, and
  2. $f(c) = m$, for some $c \in [a,b]$.

  Thus $m = \min\{f(x) : a \leq x \leq b\}$, which we also write as $\min\limits_{[a,b]} f$ or simply as $\min f$.
- f is **bounded above** on $[a,b]$ with M as an **upper bound** if
  $$f(x) \leq M, \quad \text{for all } x \in [a,b].$$
- M is the **maximum** of f on $[a,b]$ if
  1. M is an upper bound for f on $[a,b]$, and
  2. $f(d) = M$, for some $d \in [a,b]$.

  Thus $M = \max\{f(x) : a \leq x \leq b\}$, which we also write as $\max\limits_{[a,b]} f$ or simply as $\max f$.
- f is **bounded** on $[a,b]$ if it is both bounded below and bounded above on $[a,b]$.

2. Let the function f be defined on the closed interval $[a,b]$. Then

- m is the **infimum** or **greatest lower bound** of f on $[a,b]$ if
  1. m is a lower bound for f on $[a,b]$, and
  2. if $m' > m$, then $f(c) < m'$, for some $c \in [a,b]$.

  Thus $m = \inf\{f(x) : a \leq x \leq b\}$, which we also write as $\inf\limits_{[a,b]} f$ or simply as $\inf f$.
- M is the **supremum** or **least upper bound** of f on $[a,b]$ if
  1. M is an upper bound for f on $[a,b]$, and
  2. if $M' < M$, then $f(d) > M'$, for some $d \in [a,b]$.

  Thus $M = \sup\{f(x) : a \leq x \leq b\}$, which we also write as $\sup\limits_{[a,b]} f$ or simply as $\sup f$.

3. If $\min\limits_{[a,b]} f$ exists, then $\inf\limits_{[a,b]} f = \min\limits_{[a,b]} f$.

Similarly, if $\max\limits_{[a,b]} f$ exists, then $\sup\limits_{[a,b]} f = \max\limits_{[a,b]} f$.

4. A **partition** P of a closed interval $[a, b]$ is a collection of a finite number of closed subintervals of $[a, b]$,

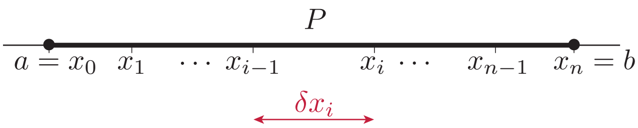
$$P = \{[x_0, x_1], \ldots, [x_{i-1}, x_i], \ldots, [x_{n-1}, x_n]\},$$

where

$$a = x_0 < x_1 < \cdots < x_i < \cdots < x_n = b.$$

The **partition points** of P are the points x_i, for $0 \le i \le n$.

The ith **subinterval** is $[x_{i-1}, x_i]$, $1 \le i \le n$, and its **length** is denoted by $\delta x_i = x_i - x_{i-1}$.



The **mesh** of P is the length of the longest subinterval: $\|P\| = \max\limits_{1 \le i \le n} \{\delta x_i\}$.

A **standard partition** is a partition with subintervals of equal length.

A **refinement** of a partition P is obtained from P by adding to it a finite number of partition points.

The **common refinement** of partitions P and P' is the partition obtained by using all their partition points.

5. Riemann sums

Let f be a bounded function on $[a, b]$, and let P be the partition

$$\{[x_0, x_1], [x_1, x_2], \ldots, [x_{i-1}, x_i], \ldots, [x_{n-1}, x_n]\},$$

where $x_0 = a$ and $x_n = b$. Let

$$m_i = \inf\{f(x) : x \in [x_{i-1}, x_i]\}$$

and

$$M_i = \sup\{f(x) : x \in [x_{i-1}, x_i]\},$$

for $i = 1, 2, \ldots, n$.

The **lower Riemann sum** for f on $[a, b]$ with partition P is

$$L(f, P) = \sum_{i=1}^{n} m_i \delta x_i.$$

The **upper Riemann sum** for f on $[a, b]$ with partition P is

$$U(f, P) = \sum_{i=1}^{n} M_i \delta x_i.$$

Theorem F43

Let f be a bounded function on $[a, b]$, and let P be a partition of $[a, b]$. Then

$$L(f, P) \le U(f, P).$$

Theorem F44

Let f be a bounded function on $[a, b]$, and let P and P' be partitions of $[a, b]$. Then

$$L(f, P) \le U(f, P').$$

6. The integral

Let f be a bounded function on $[a, b]$.

The **lower integral** of f on $[a, b]$ is

$$\underline{\int_a^b} f = \sup_P L(f, P),$$

and the **upper integral** of f on $[a, b]$ is

$$\overline{\int_a^b} f = \inf_P U(f, P),$$

where the supremum and infimum sum over all possible partitions P of the interval $[a, b]$.

We say that f is **integrable** on $[a, b]$ if

$$\underline{\int_a^b} f = \overline{\int_a^b} f.$$

The **(Riemann) integral** of f on $[a, b]$ is this common value of the lower and upper integrals.

We write $\displaystyle\int_a^b f$ or $\displaystyle\int_a^b f(x)\, dx$.

The **limits of integration** are a and b.

7. Criteria for integrability

Theorem F45

Let f be a bounded function on $[a, b]$. If there is a sequence of partitions (P_n) of $[a, b]$ such that $\|P_n\| \to 0$ and

$$\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n) = A,$$

where $A \in \mathbb{R}$, then f is integrable on $[a, b]$ and

$$\int_a^b f = A.$$

Theorem F46

If f is an integrable function on $[a, b]$ and (P_n) is a sequence of partitions of $[a, b]$ such that $\|P_n\| \to 0$, then

$$\lim_{n\to\infty} L(f, P_n) = \lim_{n\to\infty} U(f, P_n) = \int_a^b f.$$

To show that a function is not integrable, find a sequence of partitions (P_n) of $[a, b]$ whose mesh tends to zero, but for which

$$\lim_{n\to\infty} L(f, P_n) \neq \lim_{n\to\infty} U(f, P_n).$$

Corollary F47 Riemann's Criterion

Let f be bounded on $[a, b]$. Then

- f is integrable on $[a, b]$

if and only if

- there is a sequence (P_n) of partitions of $[a, b]$ with $\|P_n\| \to 0$ such that
$$U(f, P_n) - L(f, P_n) \to 0.$$

8. Two classes of integrable functions

Theorem F48

A function f which is bounded and monotonic on $[a, b]$ is integrable on $[a, b]$.

Theorem F49

A function f which is continuous on $[a, b]$ is integrable on $[a, b]$.

9. Properties of integrals

Let f be a bounded function that is integrable on an interval I containing a and b, where $a < b$. Then we make the following definitions.

- $\displaystyle\int_a^a f = 0$
- $\displaystyle\int_b^a f = -\int_a^b f$

Theorem F50 Additivity of integrals

Let f be a bounded function that is integrable on an interval I containing the points a, b and c. Then

$$\int_a^c f + \int_c^b f = \int_a^b f.$$

Theorem F51 Sign of an integral

Let f be a bounded function that is integrable on $[a, b]$.

- If $f(x) \geq 0$ on $[a, b]$, then $\displaystyle\int_a^b f \geq 0$.
- If $f(x) \leq 0$ on $[a, b]$, then $\displaystyle\int_a^b f \leq 0$.

Theorem F52 Modulus Rule

If f is integrable on $[a, b]$, then $|f|$ is also integrable on $[a, b]$.

10. Combination Rules for integrable functions

Theorem F53 Combination Rules

If f and g are integrable on $[a, b]$, then so are the following functions.

Sum Rule $f + g$, with integral
$$\int_a^b (f+g) = \int_a^b f + \int_a^b g$$

Multiple Rule λf, with integral
$$\int_a^b \lambda f = \lambda \int_a^b f \quad (\lambda \in \mathbb{R})$$

Product Rule fg

Quotient Rule f/g, provided that $1/g$ is bounded on $[a, b]$.

2 Evaluation of integrals

11. A function F is a **primitive** (or an **antiderivative**) of a function f defined on an interval I if F is differentiable on I and

$$F'(x) = f(x), \quad \text{for } x \in I.$$

A primitive F is also called an **indefinite integral** of f and denoted by

$$\int f(x)\,dx\,.$$

Integration is the process of finding a primitive of f, and the function f is called an **integrand**.

Theorem F54 Fundamental Theorem of Calculus

Let f be integrable on $[a, b]$ and let F be a primitive of f on $[a, b]$. Then

$$\int_a^b f = F(b) - F(a).$$

The expression $F(b) - F(a)$ is also written as

$$\big[F(x)\big]_a^b \quad \text{or} \quad F(x)\big|_a^b.$$

The integral

$$\int_a^b f$$

is also called the **definite integral** of f over $[a, b]$.

Theorem F55 Uniqueness Theorem for Primitives

Let F_1 and F_2 be primitives of f on an interval I. Then there exists some constant c such that

$$F_2(x) = F_1(x) + c, \quad \text{for } x \in I.$$

12. **Combination Rules for primitives**

Theorem F56 Combination Rules

Let F and G be primitives of f and g, respectively, on an interval I, and let $\lambda \in \mathbb{R}$. Then, on I:

Sum Rule $f + g$ has a primitive $F + G$

Multiple Rule λf has a primitive λF

Scaling Rule $x \longmapsto f(\lambda x)$

has a primitive

$$x \longmapsto \frac{1}{\lambda} F(\lambda x), \quad \text{for } \lambda \neq 0.$$

13. **Techniques for integration**

Strategy F8 Integration by substitution

To find a primitive $\int f(g(x))g'(x)\,dx$ using integration by substitution, do the following.

1. Choose $u = g(x)$. Find $\dfrac{du}{dx} = g'(x)$ and hence express du in terms of x and dx.
2. Substitute $u = g(x)$ and replace $g'(x)\,dx$ by du (adjusting constants if necessary) to give $\int f(u)\,du$.
3. Find $\int f(u)\,du$.
4. Substitute $u = g(x)$ to give the required primitive in terms of x.

To evaluate a definite integral, rather than find a primitive, there is no need to perform step 4 of this strategy. Instead, change the x-limits of integration into the corresponding u-limits:

$$\int_a^b f(g(x))g'(x)\,dx = \int_{g(a)}^{g(b)} f(u)\,du.$$

Also, if $g(x) > 0$ for $x \in I$, then

$$\int \frac{g'(x)}{g(x)}\,dx = \log(g(x)).$$

Strategy F9 Integration by backwards substitution

To find a primitive $\int f(x)\,dx$ using integration by backwards substitution, do the following.

1. Choose $u = g(x)$, where g has an inverse function $x = h(u)$. Find $\dfrac{dx}{du} = h'(u)$ and hence express dx in terms of u and du.
2. Substitute $x = h(u)$ and replace dx by $h'(u)\,du$ to give a primitive in terms of u.
3. Find this primitive.
4. Substitute $u = g(x)$ to give the required primitive in terms of x.

To evaluate a definite integral, rather than find a primitive, there is no need to perform step 4 of this strategy. Instead, change the x-limits of integration into the corresponding u-limits:

$$\int_a^b f(x)\,dx = \int_{g(a)}^{g(b)} f(g^{-1}(u))(g^{-1})'(u)\,du,$$

or expressed in terms of h, where $x = h(u)$,

$$\int_a^b f(x)\,dx = \int_{g(a)}^{g(b)} f(h(u))h'(u)\,du.$$

Strategy F10 Integration by parts

To find a primitive $\int k(x)\,dx$ using integration by parts, do the following.

1. Write the original function k in the form fg', where f is a function that you can differentiate and g' is a function that you can integrate.
2. Use the formula $\int fg' = fg - \int f'g$.

14. Reduction formulas

A **reduction formula** (or **recurrence relation**) involving an integral I_n (for a non-negative integer n) relates the value of I_n to the value of I_{n-1} or I_{n-2} and can be obtained by integration by parts.

3 Inequalities, sequences and series

15. Inequalities for integrals

Theorem F57 Inequality Rules

Let f and g be integrable on $[a,b]$.

(a) If $f(x) \le g(x)$, for $x \in [a,b]$, then

$$\int_a^b f \le \int_a^b g.$$

(b) If $m \le f(x) \le M$, for $x \in [a,b]$, then

$$m(b-a) \le \int_a^b f \le M(b-a).$$

Theorem F58 Triangle Inequality

Let f be integrable on $[a,b]$. Then

$$\left|\int_a^b f\right| \le \int_a^b |f|.$$

Furthermore, if $|f(x)| \le M$ for $x \in [a,b]$, then

$$\left|\int_a^b f\right| \le M(b-a).$$

16. Formulas for π

Theorem F59

(a) $$\lim_{n\to\infty}\left(\frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdot\frac{6}{7}\cdots \cdot\frac{2n}{2n-1}\cdot\frac{2n}{2n+1}\right) = \frac{\pi}{2}$$

(b) $$\lim_{n\to\infty}\frac{(n!)^2\,2^{2n}}{(2n)!\sqrt{n}} = \sqrt{\pi}$$

Wallis' Formula is the first of these formulas for π.

17. Integrals for determining convergence or divergence of series

It is often easier to evaluate an integral than a sum which has a similar behaviour, so the Integral Test can be used to help determine the convergence or divergence of certain series of the form $\sum_{n=1}^{\infty} f(n)$.

> **Theorem F60 Integral Test**
>
> Let the function f be positive and decreasing on $[1, \infty)$, and suppose that $f(x) \to 0$ as $x \to \infty$. Then
>
> (a) $\sum_{n=1}^{\infty} f(n)$ converges if the sequence
> $$\left(\int_1^n f\right)$$
> is bounded above
>
> (b) $\sum_{n=1}^{\infty} f(n)$ diverges if
> $$\int_1^n f \to \infty \text{ as } n \to \infty.$$

The number 1 in parts (a) and (b) of Theorem F60 can be replaced by any positive integer.

18. The series

$$\sum_{n=1}^{\infty} \frac{1}{n^p}$$

converges for $p > 1$, and diverges for $0 < p \leq 1$.

4 Stirling's Formula

19. Combination Rules for $\sim$

For positive functions f and g with domain $\mathbb{N}$, we write

$$f(n) \sim g(n) \quad \text{as } n \to \infty$$

to mean

$$\frac{f(n)}{g(n)} \to 1 \quad \text{as } n \to \infty.$$

Note, we often omit 'as $n \to \infty$' when writing expressions of this type.

***Warning*:** The statement

$$f(n) \sim g(n) \quad \text{as } n \to \infty$$

does *not* imply that $f(n) - g(n)$ tends to zero or is even bounded.

> **Theorem F61 Combination Rules**
>
> If $f_1(n) \sim g_1(n)$ and $f_2(n) \sim g_2(n)$, then:
>
> **Sum Rule** $\quad f_1(n)+f_2(n) \sim g_1(n)+g_2(n)$
>
> **Multiple Rule** $\quad \lambda f_1(n) \sim \lambda g_1(n)$, for $\lambda \in \mathbb{R}^+$
>
> **Product Rule** $\quad f_1(n)f_2(n) \sim g_1(n)g_2(n)$
>
> **Quotient Rule** $\quad \dfrac{f_1(n)}{f_2(n)} \sim \dfrac{g_1(n)}{g_2(n)}$.

20. Stirling's Formula

> **Theorem F62 Stirling's Formula**
>
> $$n! \sim \sqrt{2\pi n}\,(n/e)^n \quad \text{as } n \to \infty.$$

Unit F4 Power series

1 Taylor polynomials

1. Let f be differentiable on an open interval containing the point a.

The **tangent approximation** at a to f is

$$f(x) \approx f(a) + f'(a)(x-a).$$

2. Let f be n-times differentiable on an open interval containing the point a.

The **Taylor polynomial of degree n at a for f** is the polynomial

$$T_n(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

that is,

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^n.$$

The functions f and T_n have the same value at a and have equal derivatives at a for all orders up to and including n.

Taylor polynomials for successive values of n satisfy the recurrence relation

$$T_n(x) = T_{n-1}(x) + \frac{f^{(n)}(a)}{n!}(x-a)^n, \quad n \geq 1.$$

2 Taylor's Theorem

3. Approximating by a Taylor polynomial

Theorem F63 Taylor's Theorem

Let the function f be $(n+1)$-times differentiable on an open interval containing the points a and x. Then

$$\begin{aligned} f(x) &= f(a) + f'(a)(x-a) \\ &\quad + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n \\ &\quad + R_n(x) \\ &= T_n(x) + R_n(x), \end{aligned}$$

where $T_n(x)$ is the Taylor polynomial of degree n at a for f and

$$R_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1},$$

for some c between a and x.

$R_n(x)$ is the **remainder term**, or **error term** involved in approximating $f(x)$ by $T_n(x)$.

A **remainder estimate**, or **error bound**, for the approximation of $f(x)$ by $T_n(x)$, is an upper bound for $|R_n(x)| = |f(x) - T_n(x)|$.

Strategy F11 Taylor approximation at a point

To show that the Taylor polynomial T_n at a for f approximates f to a certain accuracy at a point $x \neq a$, do the following.

1. Obtain a formula for $f^{(n+1)}$.
2. Determine a number M such that
$$|f^{(n+1)}(c)| \leq M,$$
for all c between a and x.
3. Write down and simplify the remainder estimate
$$\begin{aligned} |f(x) - T_n(x)| &= |R_n(x)| \\ &\leq \frac{M}{(n+1)!}|x-a|^{n+1}. \end{aligned}$$

Strategy F12 Taylor approximation on an interval

To show that the Taylor polynomial T_n at a for f approximates f to a certain accuracy throughout an interval I of the form $[a, a+r]$, $[a-r, a]$ or $[a-r, a+r]$, where $r > 0$, do the following.

1. Obtain a formula for $f^{(n+1)}$.
2. Determine a number M such that
$$|f^{(n+1)}(c)| \leq M, \quad \text{for all } c \in I.$$
3. Write down and simplify the remainder estimate
$$\begin{aligned} |f(x) - T_n(x)| &= |R_n(x)| \\ &\leq \frac{M}{(n+1)!} r^{n+1}, \end{aligned}$$
for all $x \in I$.

Sometimes the maximum value of $|f^{(n+1)}(c)|$ can be determined, and often this is attained at an endpoint of the interval. Usually, however, any 'good enough' upper bound for $|f^{(n+1)}(c)|$ will do.

4. Taylor series

Theorem F64

Let f have derivatives of all orders on an open interval containing the points a and x. If

$R_n(x) \to 0$ as $n \to \infty$,

then

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Let f have derivatives of all orders at the point a.

The **Taylor series at a for f** is

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots.$$

The Taylor series for f is **valid** at x when it has the sum $f(x)$; that is, if $R_n(x) \to 0$ as $n \to \infty$.

The **range of validity** for a Taylor series is any set of values of x for which the Taylor series is valid.

The **sum function** of the Taylor series is the function f on any such range of validity.

Theorem F65 Basic Taylor series at 0

- $\dfrac{1}{1-x} = 1 + x + x^2 + \cdots = \displaystyle\sum_{n=0}^{\infty} x^n,$ for $|x| < 1$.
- $\sin x = x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!} - \cdots = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!},$ for $x \in \mathbb{R}$.
- $\cos x = 1 - \dfrac{x^2}{2!} + \dfrac{x^4}{4!} - \cdots = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!},$ for $x \in \mathbb{R}$.
- $e^x = 1 + x + \dfrac{x^2}{2!} + \cdots = \displaystyle\sum_{n=0}^{\infty} \frac{x^n}{n!},$ for $x \in \mathbb{R}$.
- $\log(1+x) = x - \dfrac{x^2}{2} + \dfrac{x^3}{3} - \cdots = \displaystyle\sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n},$ for $-1 < x \leq 1$.

The **alternating harmonic series** is

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots = \log 2.$$

3 Convergence of power series

5. Radius of convergence

Let $a \in \mathbb{R}$, $x \in \mathbb{R}$ and $a_n \in \mathbb{R}$, $n = 0, 1, 2, \ldots$.

A **power series at a in x**, with **coefficients** a_n is the expression

$$\sum_{n=0}^{\infty} a_n(x-a)^n = a_0 + a_1(x-a) + a_2(x-a)^2 + \cdots$$

The **centre** of the power series is a.

Theorem F66 Radius of Convergence Theorem

For a given power series $\displaystyle\sum_{n=0}^{\infty} a_n(x-a)^n$, exactly one of the following possibilities occurs.

(a) The series converges only for $x = a$.

(b) The series converges for all x.

(c) There is a number $R > 0$ such that

$$\sum_{n=0}^{\infty} a_n(x-a)^n \text{ converges if } |x-a| < R$$

and

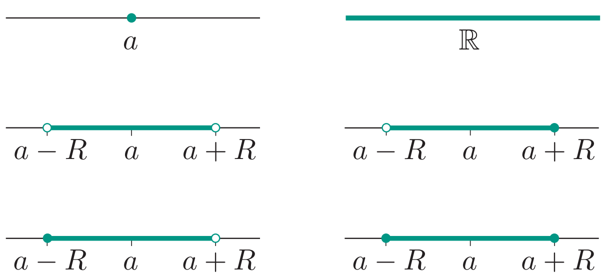
$$\sum_{n=0}^{\infty} a_n(x-a)^n \text{ diverges if } |x-a| > R.$$

Moreover, in all parts the series converges absolutely on the specified sets of convergence.

The **radius of convergence** of the power series is the positive number R in Theorem F66(c), or $R = 0$, if the power series converges only for $x = a$, or $R = \infty$, if the power series converges for all x.

***Warning*:** Theorem F66(c) makes no assertion about the behaviour of the series at the endpoints of the interval $(a - R, a + R)$.

The **interval of convergence** of the power series is the interval $(a-R, a+R)$, together with any endpoints of this interval at which the power series converges.



6. Ratio Test for power series

Theorem F67 Ratio Test

Suppose that $\sum_{n=0}^{\infty} a_n(x-a)^n$ is a power series with radius of convergence R, and that

$$\left|\frac{a_{n+1}}{a_n}\right| \to L \text{ as } n \to \infty.$$

(a) If L is ∞, then $R = 0$.
(b) If $L = 0$, then $R = \infty$.
(c) If $L > 0$, then $R = 1/L$.

Strategy F13 Finding the interval of convergence

To find the interval of convergence of the power series $\sum_{n=0}^{\infty} a_n(x-a)^n$, do the following.

1. Use the Ratio Test for power series to find the radius of convergence R.
2. If R is finite and non-zero, use other tests for series (see Strategy D13 in **15.** on page 86) to determine the behaviour of the power series at the endpoints of the interval $(a-R, a+R)$.

The largest range of validity of a Taylor series is the interval of convergence obtained by using Strategy F13.

4 Manipulating Taylor series

7. Combination Rules for Taylor series

Combination Rules

Let f and g be functions that can both be represented by a Taylor series at a, and suppose that

$$f(x) = \sum_{n=0}^{\infty} a_n(x-a)^n, \text{ for } |x-a| < R,$$

$$g(x) = \sum_{n=0}^{\infty} b_n(x-a)^n, \text{ for } |x-a| < R'.$$

The the following hold for $r = \min\{R, R'\}$ and $\lambda \in \mathbb{R}$.

Theorem F69

Sum Rule
$$(f+g)(x) = \sum_{n=0}^{\infty}(a_n + b_n)(x-a)^n \quad \text{for } |x-a| < r$$

Multiple Rule
$$\lambda f(x) = \sum_{n=0}^{\infty} \lambda a_n(x-a)^n, \quad \text{for } |x-a| < R.$$

Theorem F70

Product Rule
$$(fg)(x) = \sum_{n=0}^{\infty} c_n(x-a)^n, \quad \text{for } |x-a| < r,$$

where

$$\begin{aligned} c_n &= a_0 b_n + a_1 b_{n-1} + \cdots + a_{n-1} b_1 + a_n b_0 \\ &= \sum_{k=0}^{n} a_k b_{n-k}. \end{aligned}$$

The radius of convergence of the Taylor series for $f+g$ may be larger than $r = \min\{R, R'\}$: Theorem F69 simply asserts that it must be *at least* r.

The coefficient c_n in Theorem F70 is the result of multiplying the two Taylor series term by term and summing all the resulting coefficients of $(x-a)^n$.

8. Further rules for Taylor series

Theorem F71 Differentiation Rule

The Taylor series

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

and

$$g(x) = \sum_{n=1}^{\infty} n a_n (x-a)^{n-1}$$

have the same radius of convergence, R say.

Also, $f(x)$ is differentiable on $(a-R, a+R)$, and

$$f'(x) = g(x), \quad \text{for } |x-a| < R.$$

Theorem F72 Integration Rule

The Taylor series

$$f(x) = \sum_{n=0}^{\infty} a_n (x-a)^n$$

and

$$F(x) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x-a)^{n+1}$$

have the same radius of convergence, R say.

Also, if $R > 0$, then

$$\int f(x)\,dx = F(x), \quad \text{for } |x-a| < R.$$

The Integration Rule says that F is a primitive of f on $(a-R, a+R)$:

$$\int \left(\sum_{n=0}^{\infty} a_n (x-a)^n \right) dx = c + \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x-a)^{n+1}.$$

We find c by putting $x = a$ into this equation.

Warning: In both rules, the two series may behave differently at the endpoints of their respective intervals of convergence.

9. The **generalised binomial coefficients** are

$$\binom{\alpha}{0} = 1, \quad (\alpha \in \mathbb{R})$$

and

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!}, \quad \text{for } n \in \mathbb{N}.$$

Theorem F73 General Binomial Theorem

For $\alpha \in \mathbb{R}$,

$$(1+x)^{\alpha} = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n, \quad \text{for } |x| < 1.$$

Warning: The role of n in the generalised binomial coefficient $\binom{\alpha}{n}$ is the same as the role of k in the normal binomial coefficient $\binom{n}{k}$.

10. Techniques for finding Taylor series

- Taylor's Theorem
- the Combination Rules
- the Differentiation and Integration Rules
- the General Binomial Theorem.

Any valid method gives the same Taylor series.

Theorem F74 Uniqueness Theorem

If

$$\sum_{n=0}^{\infty} a_n (x-a)^n = \sum_{n=0}^{\infty} b_n (x-a)^n, \quad \text{for } |x-a| < R,$$

then $a_n = b_n$, for $n = 0, 1, 2, \ldots$.

5 Numerical estimates for π

11. Using the Taylor series for $\tan^{-1}$ at 0, combined with an addition formula for $\tan^{-1}$, gives the following formulas which can be used to estimate π:

$$\tan^{-1}\left(\tfrac{1}{2}\right) + \tan^{-1}\left(\tfrac{1}{3}\right) = \pi/4,$$

$$4\tan^{-1}\left(\tfrac{1}{5}\right) - \tan^{-1}\left(\tfrac{1}{239}\right) = \pi/4$$

(Machin's Formula),

$$6\tan^{-1}\left(\tfrac{1}{8}\right) + 2\tan^{-1}\left(\tfrac{1}{57}\right) + \tan^{-1}\left(\tfrac{1}{239}\right) = \pi/4.$$

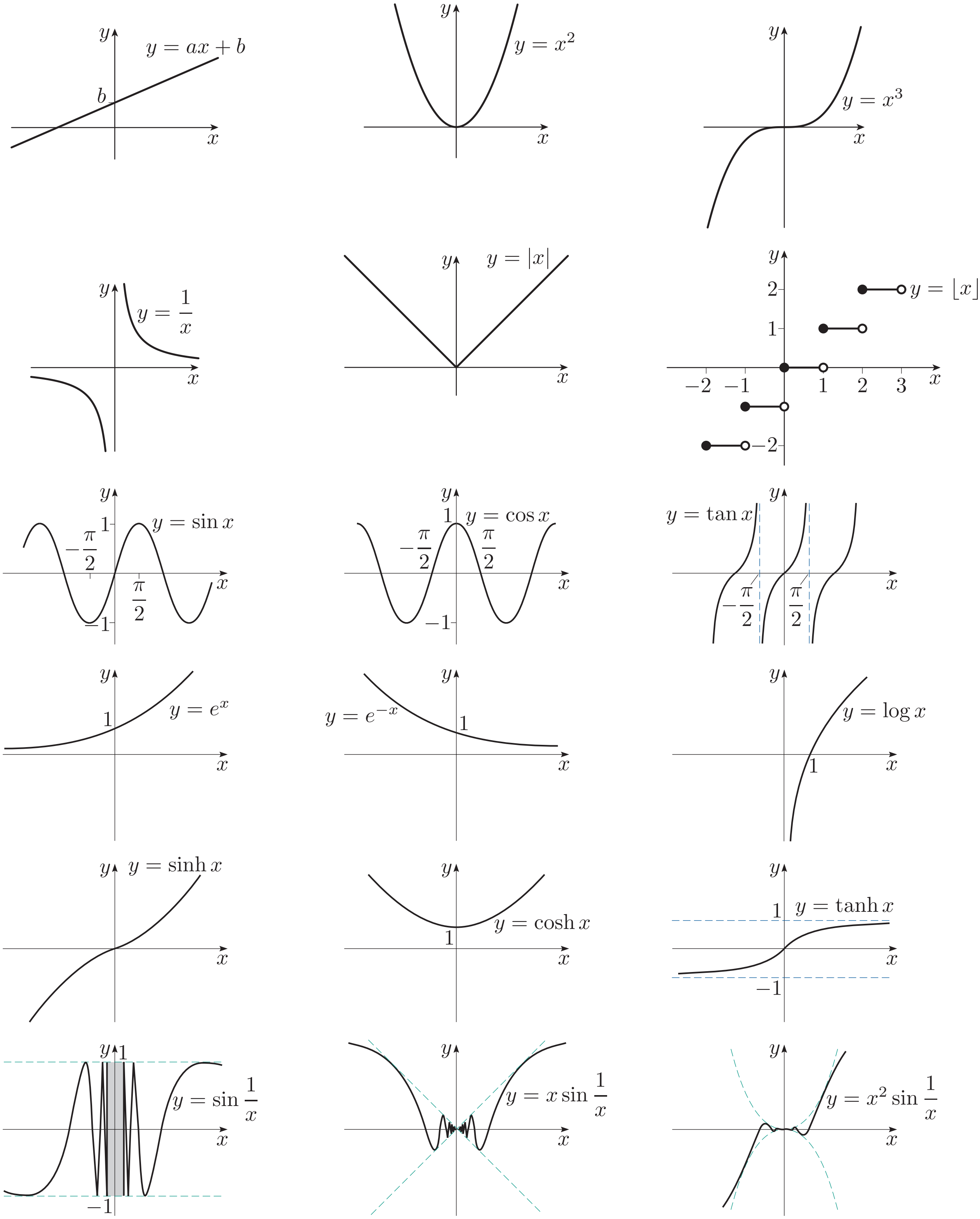
12. π is irrational

Theorem F75

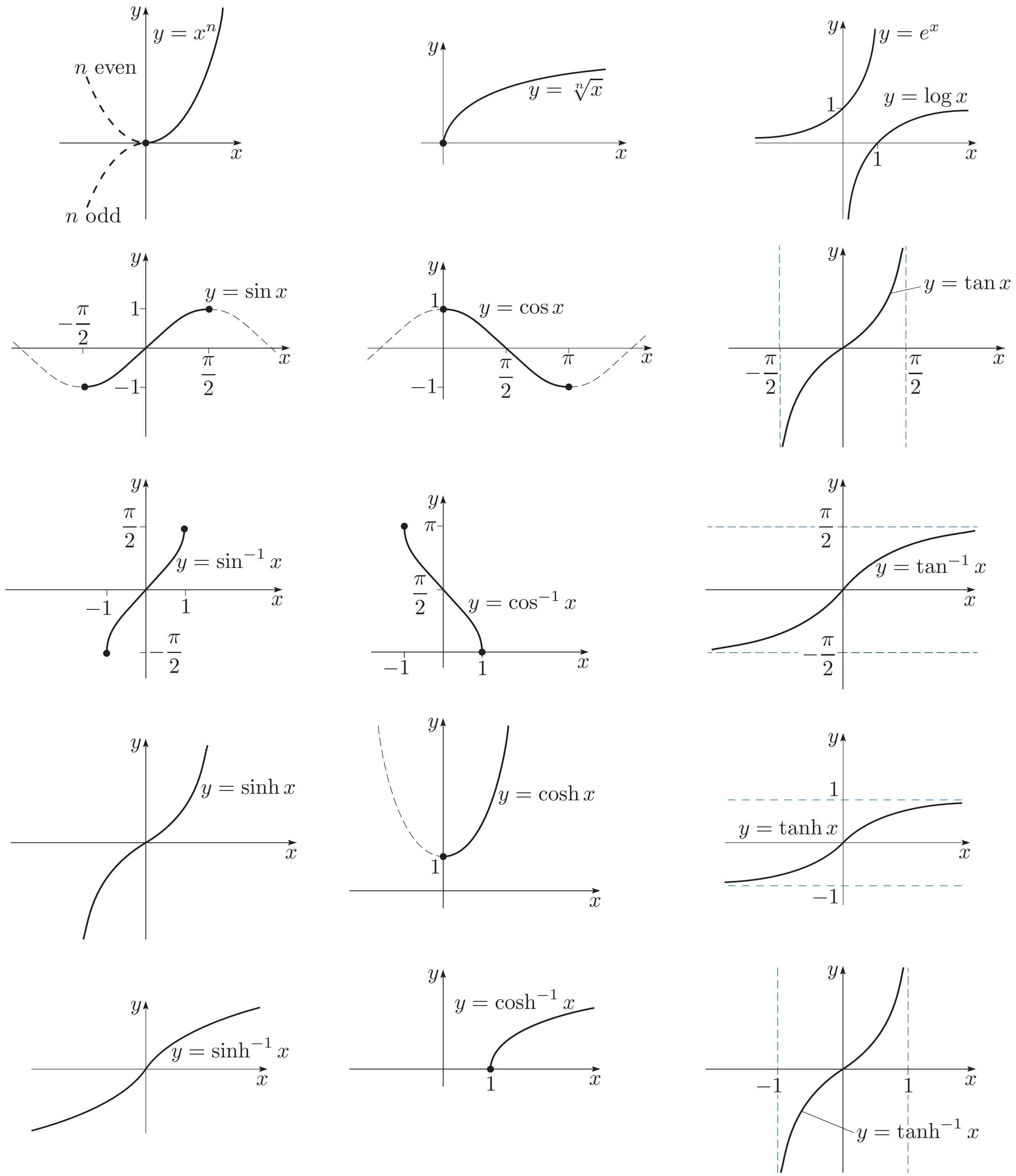
The number π is irrational.

Quick reference

Sketches of graphs of basic functions



Sketches of graphs of standard inverse functions

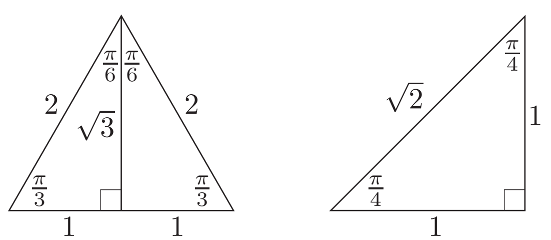


Properties of trigonometric and hyperbolic functions

Trigonometric functions	Hyperbolic functions
cos is even: $\cos(-x) = \cos x$	cosh is even: $\cosh(-x) = \cosh x$
sin is odd: $\sin(-x) = -\sin x$	sinh is odd: $\sinh(-x) = -\sinh x$
tan is odd: $\tan(-x) = -\tan x$	tanh is odd: $\tanh(-x) = -\tanh x$
$\cos^2 x + \sin^2 x = 1$	$\cosh^2 x - \sinh^2 x = 1$
$1 + \tan^2 x = \sec^2 x$	$1 - \tanh^2 x = \operatorname{sech}^2 x$
$\cot^2 x + 1 = \operatorname{cosec}^2 x$	$\coth^2 x - 1 = \operatorname{cosech}^2 x$
$\sin(x+y) = \sin x \cos y + \cos x \sin y$	$\sinh(x+y) = \sinh x \cosh y + \cosh x \sinh y$
$\cos(x+y) = \cos x \cos y - \sin x \sin y$	$\cosh(x+y) = \cosh x \cosh y + \sinh x \sinh y$
$\tan(x+y) = \dfrac{\tan x + \tan y}{1 - \tan x \tan y}$	$\tanh(x+y) = \dfrac{\tanh x + \tanh y}{1 + \tanh x \tanh y}$
$\sin 2x = 2 \sin x \cos x$	$\sinh 2x = 2 \sinh x \cosh x$
$\cos 2x = \cos^2 x - \sin^2 x$	$\cosh 2x = \cosh^2 x + \sinh^2 x$
$= 2\cos^2 x - 1$	$= 2\cosh^2 x - 1$
$= 1 - 2\sin^2 x$	$= 1 + 2\sinh^2 x$
$\tan 2x = \dfrac{2\tan x}{1 - \tan^2 x}$	$\tanh 2x = \dfrac{2\tanh x}{1 + \tanh^2 x}$
$\sin(\pi - x) = \sin x$	
$\cos(\pi - x) = -\cos x$	

Some standard values of sin, cos and tan

These can be found from the following triangles.



Standard derivatives

$f(x)$	$f'(x)$	Domain of f'
k	0	$\mathbb{R}$
x	1	$\mathbb{R}$
$x^n,\ n \in \mathbb{Z} - \{0\}$	nx^{n-1}	$\mathbb{R}$ or $\mathbb{R} - \{0\}$
$x^\alpha,\ \alpha \in \mathbb{R}$	$\alpha x^{\alpha-1}$	$\mathbb{R}^+$
$a^x,\ a > 0$	$a^x \log a$	$\mathbb{R}$
$\sin x$	$\cos x$	$\mathbb{R}$
$\cos x$	$-\sin x$	$\mathbb{R}$
$\tan x$	$\sec^2 x$	$\mathbb{R} - \left\{\left(n+\frac{1}{2}\right)\pi : n \in \mathbb{Z}\right\}$
$\operatorname{cosec} x$	$-\operatorname{cosec} x \cot x$	$\mathbb{R} - \{n\pi : n \in \mathbb{Z}\}$
$\sec x$	$\sec x \tan x$	$\mathbb{R} - \left\{\left(n+\frac{1}{2}\right)\pi : n \in \mathbb{Z}\right\}$
$\cot x$	$-\operatorname{cosec}^2 x$	$\mathbb{R} - \{n\pi : n \in \mathbb{Z}\}$
$\sin^{-1} x$	$1/\sqrt{1-x^2}$	$(-1, 1)$
$\cos^{-1} x$	$-1/\sqrt{1-x^2}$	$(-1, 1)$
$\tan^{-1} x$	$1/\left(1+x^2\right)$	$\mathbb{R}$
e^x	e^x	$\mathbb{R}$
$\log x$	$1/x$	$\mathbb{R}^+$
$\sinh x$	$\cosh x$	$\mathbb{R}$
$\cosh x$	$\sinh x$	$\mathbb{R}$
$\tanh x$	$\operatorname{sech}^2 x$	$\mathbb{R}$
$\sinh^{-1} x$	$1/\sqrt{1+x^2}$	$\mathbb{R}$
$\cosh^{-1} x$	$1/\sqrt{x^2-1}$	$(1, \infty)$
$\tanh^{-1} x$	$1/\left(1-x^2\right)$	$(-1, 1)$

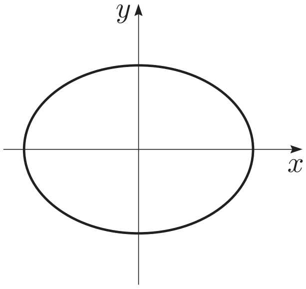
Standard primitives

$f(x)$	Primitive $F(x)$	Domain
$x^n,\ n \in \mathbb{Z} - \{-1\}$	$x^{n+1}/(n+1)$	$\mathbb{R}$ or $\mathbb{R} - \{0\}$
$x^\alpha,\ \alpha \neq -1$	$x^{\alpha+1}/(\alpha+1)$	$\mathbb{R}^+$
$a^x,\ a > 0$	$a^x/\log a$	$\mathbb{R}$
$\sin x$	$-\cos x$	$\mathbb{R}$
$\cos x$	$\sin x$	$\mathbb{R}$
$\tan x$	$\log(\sec x)$	$\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$
e^x	e^x	$\mathbb{R}$
$1/x$	$\log x$	$(0, \infty)$
$1/x$	$\log \lvert x\rvert$	$(-\infty, 0)$
$\log x$	$x \log x - x$	$(0, \infty)$
$\sinh x$	$\cosh x$	$\mathbb{R}$
$\cosh x$	$\sinh x$	$\mathbb{R}$
$\tanh x$	$\log(\cosh x)$	$\mathbb{R}$
$(a^2 - x^2)^{-1},\ a \neq 0$	$\dfrac{1}{2a} \log\left(\dfrac{a+x}{a-x}\right)$	$(-a, a)$
$(a^2 + x^2)^{-1},\ a \neq 0$	$\dfrac{1}{a} \tan^{-1}(x/a)$	$\mathbb{R}$
$(a^2 - x^2)^{-1/2},\ a \neq 0$	$\begin{cases} \sin^{-1}(x/a) \\ -\cos^{-1}(x/a) \end{cases}$	$(-a, a)$ $(-a, a)$
$(x^2 - a^2)^{-1/2},\ a \neq 0$	$\begin{cases} \log(x + (x^2 - a^2)^{1/2}) \\ \cosh^{-1}(x/a) \end{cases}$	(a, ∞) (a, ∞)
$(a^2 + x^2)^{-1/2},\ a \neq 0$	$\begin{cases} \log(x + (a^2 + x^2)^{1/2}) \\ \sinh^{-1}(x/a) \end{cases}$	$\mathbb{R}$ $\mathbb{R}$
$(a^2 - x^2)^{1/2},\ a \neq 0$	$\frac{1}{2}x(a^2 - x^2)^{1/2} + \frac{1}{2}a^2 \sin^{-1}(x/a)$	$(-a, a)$
$(x^2 - a^2)^{1/2},\ a \neq 0$	$\frac{1}{2}x(x^2 - a^2)^{1/2} - \frac{1}{2}a^2 \log(x + (x^2 - a^2)^{1/2})$	(a, ∞)
$(a^2 + x^2)^{1/2},\ a \neq 0$	$\frac{1}{2}x(a^2 + x^2)^{1/2} + \frac{1}{2}a^2 \log(x + (a^2 + x^2)^{1/2})$	$\mathbb{R}$
$e^{ax} \cos bx,\ a, b \neq 0$	$\dfrac{e^{ax}}{a^2 + b^2}(a \cos bx + b \sin bx)$	$\mathbb{R}$
$e^{ax} \sin bx,\ a, b \neq 0$	$\dfrac{e^{ax}}{a^2 + b^2}(a \sin bx - b \cos bx)$	$\mathbb{R}$

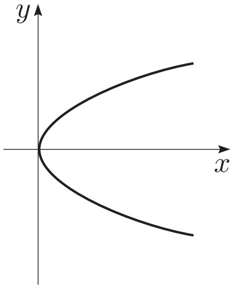
Standard Taylor series

Function	Taylor series	Domain
$\dfrac{1}{1-x}$	$1+x+x^2+x^3+\cdots=\displaystyle\sum_{n=0}^{\infty}x^n$	$\lvert x\rvert<1$
$\log(1+x)$	$x-\dfrac{x^2}{2}+\dfrac{x^3}{3}-\dfrac{x^4}{4}+\cdots=\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n+1}x^n}{n}$	$-1<x\leq 1$
e^x	$1+x+\dfrac{x^2}{2!}+\dfrac{x^3}{3!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{x^n}{n!}$	$x\in\mathbb{R}$
$\sin x$	$x-\dfrac{x^3}{3!}+\dfrac{x^5}{5!}-\dfrac{x^7}{7!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{(2n+1)!}$	$x\in\mathbb{R}$
$\cos x$	$1-\dfrac{x^2}{2!}+\dfrac{x^4}{4!}-\dfrac{x^6}{6!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n}}{(2n)!}$	$x\in\mathbb{R}$
$(1+x)^\alpha$	$1+\alpha x+\dfrac{\alpha(\alpha-1)}{2!}x^2+\dfrac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\cdots=\displaystyle\sum_{n=0}^{\infty}\binom{\alpha}{n}x^n$	$\lvert x\rvert<1,\ \alpha\in\mathbb{R}$
$\sinh x$	$x+\dfrac{x^3}{3!}+\dfrac{x^5}{5!}+\dfrac{x^7}{7!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{x^{2n+1}}{(2n+1)!}$	$x\in\mathbb{R}$
$\cosh x$	$1+\dfrac{x^2}{2!}+\dfrac{x^4}{4!}+\dfrac{x^6}{6!}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{x^{2n}}{(2n)!}$	$x\in\mathbb{R}$
$\tan^{-1}x$	$x-\dfrac{x^3}{3}+\dfrac{x^5}{5}-\dfrac{x^7}{7}+\cdots=\displaystyle\sum_{n=0}^{\infty}\frac{(-1)^n x^{2n+1}}{2n+1}$	$\lvert x\rvert\leq 1$

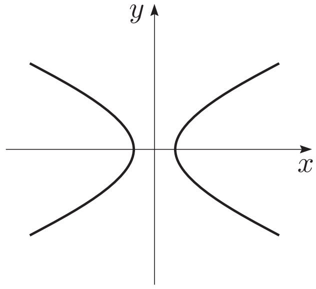
Non-degenerate conics in standard position



Ellipse $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$



Parabola $y^2 = 4ax$

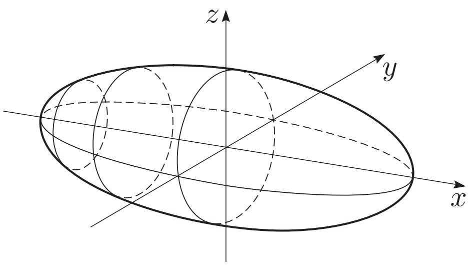


Hyperbola $\dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1$

Quadrics in standard position – the six types considered

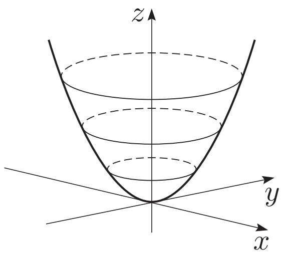
The curves of intersection are given below each quadric.

Ellipsoid $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} + \dfrac{z^2}{c^2} = 1$



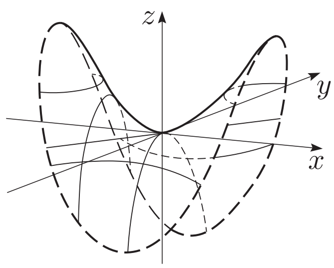
ellipse

Elliptic paraboloid $z = \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2}$



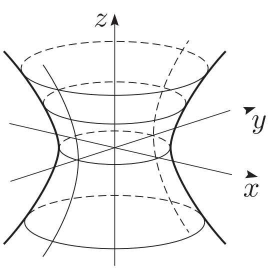
ellipse or parabola

Hyperbolic paraboloid $z = \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2}$



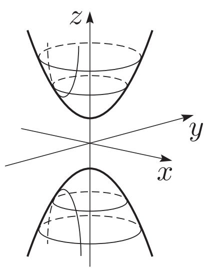
hyperbola or parabola

Hyperboloid of one sheet $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 1$



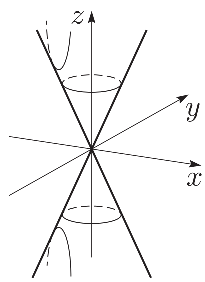
ellipse or hyperbola

Hyperboloid of two sheets $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = -1$



ellipse or hyperbola

Elliptic cone $\dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} - \dfrac{z^2}{c^2} = 0$



ellipse or hyperbola
(or a degenerate conic)

Standard groups

Throughout, $m, n \in \mathbb{N}$ and p is prime.

Finite groups of small order

Group(s)	Order
$C_2 \cong (\mathbb{Z}_2, +_2)$	2
$C_3 \cong (\mathbb{Z}_3, +_3)$	3
$C_4 \cong (\mathbb{Z}_4, +_4)$	4
$V \cong S(\square)$	4

Group(s)	Order
$C_5 \cong (\mathbb{Z}_5, +_5)$	5
$C_6 \cong (\mathbb{Z}_6, +_6)$	6
$S(\triangle) \cong D_3$	6
$C_7 \cong (\mathbb{Z}_7, +_7)$	7

Group(s)	Order
$C_8 \cong (\mathbb{Z}_8, +_8)$	8
$S(\text{cuboid})$	8
$(U_{15}, \times_{15})$	8
$S(\square) \cong D_4$	8
Q_8	8

Group(s)	Order
A_4	12
S_4	24

Finite groups – infinite families

Group	Order	Description
C_n	n	standard abstract cyclic group (multiplicative)
$(\mathbb{Z}_n, +_n)$	n	standard cyclic group (additive), $n \geq 2$
$(\mathbb{Z}_p^*, \times_p)$	$p-1$	an abelian group: $\mathbb{Z}_p^* = \{1, 2, \ldots, p-1\}$
$(U_n, \times_n)$	–	an abelian group: U_n is the set of all integers in $\mathbb{Z}_n$ coprime to n, $n \geq 2$
S_n	$n!$	the symmetric group of degree n
A_n	$n!/2$	the alternating group of degree n, $n \geq 2$
D_n	$2n$	the dihedral group of order $2n$ (the symmetries of an n-gon), $n \geq 3$

Infinite groups (other than groups of matrices)

$(\mathbb{Z}, +)$	$(\mathbb{Q}, +)$	$(\mathbb{R}, +)$	$(\mathbb{C}, +)$	$(\mathbb{R}^2, +)$	$(\mathbb{R}^3, +)$	$(\mathbb{R}^n, +)$	$(\mathbb{R}^\infty, +)$
	$(\mathbb{Q}^*, \times)$	$(\mathbb{R}^*, \times)$	$(\mathbb{C}^*, \times)$				
	$(\mathbb{Q}^+, \times)$	$(\mathbb{R}^+, \times)$					

Infinite groups of matrices

Group	Set
$(M_{2,2}, +)$	2×2 matrices
$(\mathrm{GL}(2), \times)$	invertible 2×2 matrices
$(\mathrm{SL}(2), \times)$	2×2 matrices with determinant 1
$(D, \times)$	invertible 2×2 diagonal matrices
$(L, \times)$	invertible 2×2 lower triangular matrices
$(U, \times)$	invertible 2×2 upper triangular matrices
$(M_{m,n}, +)$	$m \times n$ matrices
$(\mathrm{GL}(n), \times)$	invertible $n \times n$ matrices
$(\mathrm{SL}(n), \times)$	$n \times n$ matrices with determinant 1

$$\mathrm{GL}(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R},\ ad - bc \neq 0 \right\}$$

$$\mathrm{SL}(2) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{R},\ ad - bc = 1 \right\}$$

$$D = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} : a, d \in \mathbb{R},\ ad \neq 0 \right\}$$

$$L = \left\{ \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} : a, c, d \in \mathbb{R},\ ad \neq 0 \right\}$$

$$U = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} : a, b, d \in \mathbb{R},\ ad \neq 0 \right\}$$

Isomorphism classes for groups of orders 1 to 8

Isomorphism classes for groups of orders 1 to 7

Order	Standard group(s)	Distinguishing features (given the order of the group)	Further examples
1	C_1		$(\{0\}, +)$, $(\{1\}, \times)$
2	C_2, $\mathbb{Z}_2$		$S^+(\text{▭})$, $(\mathbb{Z}_3^*, \times_3)$
3	C_3, $\mathbb{Z}_3$		$S^+(\triangle)$, $(\{0, 4, 8\}, +_{12})$, $(\{1, 4, 7\}, \times_9)$
4	C_4, $\mathbb{Z}_4$	Exactly 2 self-inverse elements. An element of order 4.	$(\mathbb{Z}_5^*, \times_5)$, $S^+(\square)$, $S(\text{✤})$, $(\{0, 3, 6, 9\}, +_{12})$, $(\{1, -1, i, -i\}, \times)$
	V, $S(\text{▭})$	All elements self-inverse. (*Note that V, like C_4, is abelian.*)	$(U_8, \times_8)$, $(U_{12}, \times_{12})$, $(\{1, 7, 9, 15\}, \times_{16})$, $(\{1, 9, 11, 19\}, \times_{20})$
5	C_5, $\mathbb{Z}_5$		$S^+(\text{⬠})$
6	C_6, $\mathbb{Z}_6$	Abelian. An element of order 6.	$S^+(\text{⬡})$, $(\mathbb{Z}_7^*, \times_7)$, $(U_9, \times_9)$, $(U_{14}, \times_{14})$ $(\{0, 2, 4, 6, 8, 10\}, +_{12})$
	$S(\triangle)$	Non-abelian.	S_3, $\{e, (2\,3), (2\,4), (3\,4), (2\,3\,4), (2\,4\,3)\}$
7	C_7, $\mathbb{Z}_7$		$S^+(\text{heptagon})$

Where two sets of distinguishing features are given on separate lines in the same row of the table, either *distinguishes the isomorphism class.*

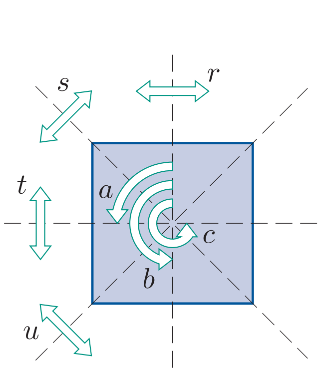
Proposition B80 Isomorphism classes for groups of order 8

There are five isomorphism classes for groups of order 8, as follows.

Class	Abelian/ non-abelian	Numbers of elements of order 1	order 2	order 4	order 8	Example
1	abelian	1	1	2	4	$(\mathbb{Z}_8, +_8)$
2	abelian	1	7	0	0	$S(\text{cuboid})$
3	abelian	1	3	4	0	$(U_{15}, \times_{15})$
4	non-abelian	1	5	2	0	$S(\square)$
5	non-abelian	1	1	6	0	Q_8

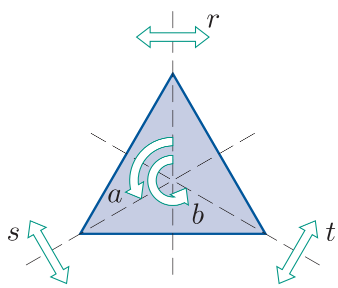
Group tables of symmetry groups

$S(\square)$



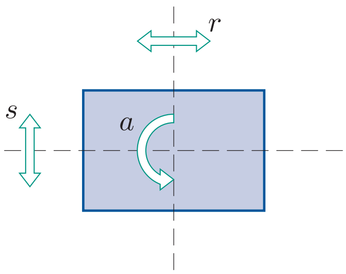
$\circ$	e	a	b	c	r	s	t	u
e	e	a	b	c	r	s	t	u
a	a	b	c	e	s	t	u	r
b	b	c	e	a	t	u	r	s
c	c	e	a	b	u	r	s	t
r	r	u	t	s	e	c	b	a
s	s	r	u	t	a	e	c	b
t	t	s	r	u	b	a	e	c
u	u	t	s	r	c	b	a	e

$S(\triangle)$



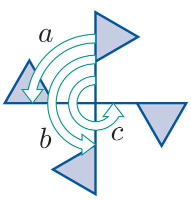
$\circ$	e	a	b	r	s	t
e	e	a	b	r	s	t
a	a	b	e	t	r	s
b	b	e	a	s	t	r
r	r	s	t	e	a	b
s	s	t	r	b	e	a
t	t	r	s	a	b	e

$S(▭) \cong V$



$\circ$	e	a	r	s
e	e	a	r	s
a	a	e	s	r
r	r	s	e	a
s	s	r	a	e

$S(✢) \cong C_4$



$\circ$	e	a	b	c
e	e	a	b	c
a	a	b	c	e
b	b	c	e	a
c	c	e	a	b

Complex numbers: Cartesian, polar and exponential form

Let $z = x + iy = r(\cos\theta + i\sin\theta) = re^{i\theta}$,

$$z_1 = x_1 + iy_1 = r_1(\cos\theta_1 + i\sin\theta_1) = r_1e^{i\theta_1},$$
$$z_2 = x_2 + iy_2 = r_2(\cos\theta_2 + i\sin\theta_2) = r_2e^{i\theta_2}.$$

Complex conjugate

Cartesian form $\overline{z} = x - iy$

Polar form $\overline{z} = r(\cos\theta + i\sin(-\theta))$
$(= r(\cos\theta - i\sin\theta))$

Exponential form $\overline{z} = re^{-i\theta}$

Product

Cartesian form Use the usual rules of arithmetic to find z_1z_2 (with $i^2 = -1$).

Polar form
$$z_1z_2 = r_1r_2\left(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right)$$

Exponential form $z_1z_2 = r_1r_2e^{i(\theta_1+\theta_2)}$

Reciprocal (In each case, $z \neq 0$.)

Cartesian form $\dfrac{1}{z} = \dfrac{1}{z} \times \dfrac{\overline{z}}{\overline{z}} = \dfrac{\overline{z}}{|z|^2}$

Polar form $\dfrac{1}{z} = \dfrac{1}{r}(\cos(-\theta) + i\sin(-\theta))$
$\left(= \dfrac{1}{r}(\cos\theta - i\sin\theta)\right)$

Exponential form $\dfrac{1}{z} = \dfrac{1}{r}e^{-i\theta}$

Quotient (In each case, $z_2 \neq 0$.)

Cartesian form $\dfrac{z_1}{z_2} = \dfrac{z_1}{z_2} \times \dfrac{\overline{z_2}}{\overline{z_2}} = \dfrac{z_1\overline{z_2}}{|z_2|^2}$

Polar form $\dfrac{z_1}{z_2} = \dfrac{r_1}{r_2}(\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2))$

Exponential form $\dfrac{z_1}{z_2} = \dfrac{r_1}{r_2}e^{i(\theta_1-\theta_2)}$

Converting polar and exponential form to Cartesian form

Use the equations

$$x = r\cos\theta, \quad y = r\sin\theta.$$

Converting Cartesian form to polar and exponential form

Find the modulus r, using $r = |z| = \sqrt{x^2+y^2}$.

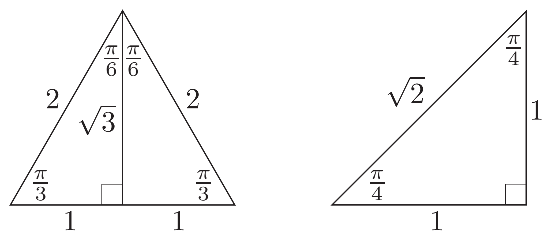
Mark z on a sketch of the complex plane. Find the acute angle ϕ at the origin in the right-angled triangle formed by drawing the perpendicular from z to the real axis, using

$$\cos\phi = \frac{|x|}{r}.$$

Hence find the principal argument.

Sines and cosines of special angles

θ	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$
$\sin\theta$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1
$\cos\theta$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0



Triangles for finding sines and cosines of special angles

Useful trigonometric identities

For any $\theta \in \mathbb{R}$,

$$\sin(\pi-\theta) = \sin\theta, \qquad \sin(-\theta) = -\sin\theta,$$
$$\cos(\pi-\theta) = -\cos\theta, \qquad \cos(-\theta) = \cos\theta.$$

For any $\theta_1, \theta_2 \in \mathbb{R}$,

$$\sin(\theta_1+\theta_2) = \sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2,$$
$$\cos(\theta_1+\theta_2) = \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2.$$

Index of strategies

Index

Index entries referring to unit summaries are in the form ‘page number(item number)’.